University of California, Berkeley
Department of Music

The 1999 Ernest Bloch Lectures

Music and Mind:

Foundations of Cognitive Musicology

David Huron

The 1999 Ernest Bloch Lectures endeavor to introduce the field of cognitive musicology to a general audience. The lectures address questions concerning the origins of music, the emotional experience of music, the relationship between music and culture, and questions of musical taste and value. Professor Huron argues that investigating the musical mind is one of the central tasks of music scholarship.

Six lectures are scheduled:

Lecture 1: Music and Mind

> "My first lecture provides an introductory tour of the field of cognitive musicology. The lecture traces some of the history of the field, clarifies some of the premises and assumptions that motivate scholars, and relays some sample research accomplishments in the areas of music performance, composition, perception, music history, and in social and cultural areas. My hope is that this first lecture will convey some of the flavor for what people do in the discipline, and why it might matter to other music scholars, musicians, and music-lovers."

Lecture 2: Is Music an Evolutionary Adaptation?

"My second lecture addresses the question of music's origins. The archeological evidence suggests that music is at least 50,000 years old, and perhaps a quarter of a million years old. In addition to the archeological evidence, there is biochemcial, neurological, behavioral and anthropological evidence that suggests that it is possible that music (or aspects of music) may be an evolutionary adaptation. The motivation for Lecture 2 is not to convince my audience that there are genes for music. Rather, what I hope to do is convince my audience that the evidence for music as an evolutionary adaptation is at least as strong as comparable evidence that has been advanced supporting the idea that language is an evolutionary adaptation."

Lecture 3: Methodology

"Cognitive musicology lies at the intersection between the sciences and the arts. It is an intersection that has produced many head-on collisions between scientific and humanities approaches to scholarly research. In particular, cognitive musicology directly faces the methodological schism between *empiricism* and *post-modernism*. In my third lecture on methodology, I will attempt to explain contemporary empiricism to humanities scholars and to explain post-modernism to scientists. I will then re-interpret both of these methodological currents in a way that shows they are different sides of the same coin we call skepticism. I will also attempt to identify the circumstances when a scholar should choose one or another method in the course of their investigations."

Lecture 4: What is a Musical Feature?

"In talking about musical works, musical styles, and musical cultures, it is essential to consider the descriptive languages we use. In my fourth lecture, I will ask the question "What is a musical feature?" I will illustrate my lecture by analyzing a movement from the first string quartet by Johannes Brahms. I will contrast my analysis with a well-known set-theoretic analysis done by Professor Allan Forte, and will show how a set-theoretic analysis fails to capture musically important features. The motivation for this analysis is not to discredit Prof. Forte. Rather, the motivation is to establish some criteria that lead to greater clarity in how we describe artifacts. That is, Lecture 4 will address the question of how to evaluate a music analysis."

Lecture 5: A Theory of Music and Affect

"In my fifth lecture, I will delve into the area of music and emotions. The emotional dimension of musical experience has been poorly served by conventional music scholarship. But it is an area of investigation that is especially well served by a cognitive approach. In this lecture I will present a theory of how music evokes emotions."

This lecture is divided into three "chapters." Only Chapter 1 ("Musical Expectation") is currently available online.

Lecture 6: A Cognitive Anthropology for Music

"In the sixth and final lecture, I will examine how a cognitive approach can illuminate the social and cultural bases of music. Drawing on the field of cognitive anthropology, I will give a number of examples of research projects -- many of which I've been involved in -- that examine cultural differences and similarities from a cognitive perspective."

University of California, Berkeley
Department of Music

The 1999 Ernest Bloch Lectures

Music and Mind:

Foundations of Cognitive Musicology

David Huron

Preface

The field of music cognition has seen a dramatic increase in activity in the past decade. It is the tendency of scholars to hold the belief that their own field of endeavor is somehow central to the enterprise of human knowledge. This belief can be found among scholars of all persuasions. What the dedicated historian, the philosopher, the medical practitioner, the chemist, the therapist and the artist all share is a belief in the importance of their field of work.

A sense of self-importance is not merely a delusion, although there is surely an element of delusion present. Those who are immersed daily in the minutiae of a field of study are typically those who are most able to see the extent, the importance, and (possibly) the grandeur of the collective project in which they and their colleagues are engaged. A sense of importance also serves a practical purpose. For what better way to motivate hard work than to hold the conviction that what one does matters?

In a lecture series such as this, the members of the audience have certain expectations, and the speaker has certain obligations. There exists a well-defined rhetorical schema where the lecturer will engage in outlining a grand design, and will demonstrate how her or his field of activity and methodological approach hold significant meaning or promise significant new insights. We expect the lecturer to make bold yet circumspect new claims, to identify widespread misconceptions, to tell a few good stories, and to point to a promised land of scholarly wisdom.

Lest you fear that I have resolved not to deliver the goods, let me reassure you. I am just as deluded as the next scholar, and do indeed feel that my chosen field, cognitive musicology, is, in the grand scheme of things, somehow important. I will indeed indulge in some bold (if circumspect) claims, identify some widespread misconceptions, tell a few stories, and point to what I think is a promised land of greater musical insight. In short, my lectures will provide an apology for the field of cognitive musicology.

Outline of Lectures

For those of you with the stamina to hear all six lectures, let me outline what I plan to do. In this evening's lecture I propose to provide an introductory tour of the field; to trace some of the history, make clear some of the premises and assumptions that motivate scholars; and to relay some sample research accomplishments in music performance, composition, perception, music history, and in social and cultural areas. My hope is that this first lecture will convey some of the flavor for what people do in the discipline, and why it might matter to other music scholars, musicians, and music-lovers.

In the second lecture I will address the question of music's origins. The archeological evidence suggests that music is at least 50,000 years old, and perhaps a quarter of a million years old. In addition to the archeological evidence, there is biochemical, neurological, behavioral and anthropological evidence that suggests that it is possible that music (or aspects of music) may be an evolutionary adaptation. The motivation for Lecture 2 is not to convince you that there are genes for music. Rather, what I hope to do is show that the evidence for music as an evolutionary adaptation is at least as strong as comparable evidence that has been advanced supporting the idea that language is an evolutionary adaptation.

Cognitive musicology lies at the intersection between the sciences and the arts. By "intersection" I have less in mind the idealized geometric point envisioned by Euclid, and more the sort of intersection that we see at the corner of Sacramento and University streets. The sciences and humanities have been entangled in a number of head-on collisions -- most recently, in the methodological quarrel between empiricism and post-modernism. In my third lecture, on methodology, I will attempt to explain contemporary empiricism to humanities scholars and to explain post-modernism to scientists. I will then re-interpret both of these methodological currents in a way that shows they are different sides of the coin we call skepticism. I will also attempt to identify the circumstances when scholars should choose one or another method in the course of their investigations.

In talking about musical works, musical styles, and musical cultures, it is essential to consider the descriptive languages we use. In my fourth lecture, I will address the question "What is a Musical Feature?" I will illustrate my lecture by analyzing a movement from the first string quartet by Johannes Brahms. I will contrast my analysis with a well-known set-theoretic analysis done by Professor Allan Forte, and will show how a set-theoretic analysis fails to capture musically important features. The motivation for this analysis is not to discredit Prof. Forte. Rather, the motivation is to establish some criteria that lead to greater clarity in how we describe musical works and repertories. That is, Lecture 4 will consider how we should go about evaluating a music analysis.

In my fifth lecture, I will delve into the area of music and emotions. The emotional dimension of musical experience has been poorly treated by conventional music scholarship. But it is an area of investigation that lends itself well to a cognitive approach. In this lecture I will present a theory of how music evokes emotions.

In the sixth and final lecture, I will examine how a cognitive approach can illuminate the social and cultural bases of music. Drawing on the field of cognitive anthropology, I will give a number of examples of research projects -- many of which I've been involved in -- that examine cultural differences and similarities from a cognitive perspective.

University of California, Berkeley
Department of Music

The 1999 Ernest Bloch Lectures

Lecture 1. Music and Mind:

Foundations of Cognitive Musicology

David Huron

The Origins of Cognitive Musicology

In an introductory lecture such as this, I suppose a good place to start is to address three questions: What is cognitive musicology? How did the field arise? What does it hope to achieve? Let me begin first with a thumbnail history of the origins of cognitive musicology, and then draw on this background to identify what I think are the defining features of the field. Of course practitioners in a field are rarely the best historians; so I approach the idea of tracing origins of cognitive musicology with trepidation. At the same time, I believe that reviewing some of the history can prove informative in understanding how and why the field has developed as it has.

Cognitive musicology has its origins in two intellectual currents. The first is the so-called "cognitive revolution" and the second is what might be called "music psychology." The cognitive revolution is a broad movement that has transformed psychology over the past three decades. Many music scholars with an interest in psychology have simply been swept along the path of the cognitive revolution. At the same time cognitive musicology can also be viewed an off-shoot from a century-old research tradition of music psychology -- a field whose predominantly German origins recommends using the designation *Psychologie der Musik*. However, cognitive musicology arose, at least in part, in response to specific criticism of the practice of music psychology. Please indulge me while I attempt to trace these two converging histories of scholarship.

From Music Psychology to Cognitive Musicology

There are many interesting questions one can ask about music. Why are some people more musical than others? Is musical "intelligence" independent of general intelligence? How does music give pleasure? Why do people disagree about musical likes and dislikes? Are musical preferences related to personality? Why do our musical preferences sometimes change over time? Does everyone "hear" music the same way? With training, how might we listen differently? Are there certain life experiences (such as ecstasy or grief) that contribute to a person's understanding of music? Is music somehow similar to speech or language? What makes something sound "musical?" Why do some melodies get stuck in your head? Why don't all melodies get stuck in your head? Why do people willingly listen to music that makes them sad? Can music somehow corrupt or enhance moral behavior? Can a person listen to too much music? Can we hear/understand the music of another culture in the

same way as people from that culture do? Why do cultures or styles change? Does a music tell us something about the people who make it? Can one musical culture ever be regarded as superior to another culture? What is the relationship between music and the other arts? Are there limits to what music could be?

Most of these questions are essentially psychological in nature. For the non-professional, these look like good questions -- the sort of questions that would animate music scholars. Yet professionals know that most music scholarship flits around the periphery of such questions. Unfortunately, despite a history of research going back at least 150 years, music psychology never really captured the imaginations of music scholars and so failed to become a core discipline within 20th-century musicology. There are reasons for this. Some 50 years ago, Paul Farnsworth gave a lecture on this very campus outlining what he considered the main shortcomings of music psychology. His talk was entitled "Sacred Cows in the Psychology of Music." Although I disagree with some points raised by Farnsworth, a half century later I find myself extending and refining Farnsworth's criticisms of the continuing field of music psychology. There are, I believe, at least four problems that have haunted music psychology.

1. First, throughout its history, music psychology has tended to focus on the individual, and on individual responses to music. Music psychologists often pay little attention to social and cultural context. Although early sociologists like Max Weber wrote extensively about music, later social psychologists failed to continue the tradition.[1]
2. Secondly, although psychology is a broad discipline, music psychology has tended to focus exclusively on low-level issues of sensation and perception. While many significant discoveries have been made, these discoveries have held little pertinence to musical experience. To this day, most books on the psychology of music typically include lengthy discussions of acoustics and psychophysics without showing how these matters might relate to the quality of musical experience.
3. Thirdly, when music psychology has addressed more musically interesting questions -- such as (say) the perceptibility of serial transformations -- the resulting research has tended to emphasize the limitations of music listening. Again and again, music psychologists have been the bearers of bad news. All of this nay-saying might have been offset if music psychologists had shown a comparable interest in discussing what music *might be*. That is, the discipline has lacked a creative or imaginative component; in general, it has not spawned research ventures that point to new and unexplored musical terrain. Until very recently, a composer reading works in the psychology of music would find little inspiration.
4. Finally, the field of music psychology has tended to be dominated by researchers with conservative musical tastes. Well-known researchers like Carl Seashore have shown little interest in contemporary music, and many mid-twentieth century music psychologists were privately or openly hostile towards the new music. Practicing musicians have been largely justified in suspecting music psychologists of pursuing a conservative musical agenda. It should be noted that the discipline itself has attracted scholars who are suspicious of the new music, and who think that psychological research can be used to buttress their arguments that contemporary music is somehow "unnatural."

To be fair to my colleagues and predecessors, one needs to include some rejoinders to these four criticisms.

1. First, in carrying out any research program, one must narrow the field of inquiry if anything is to be accomplished. The topic on which one focuses often arises from convenience. (If a music theorist chooses to analyse a particular work, it does not necessarily follow that the theorist thinks other works unworthy of study.) Music psychologists focused on individual responses rather than broader social and cultural issues primarily because it is easier to study individuals rather than groups.
2. Secondly, the emphasis on low-level aspects of sensation and perception has proved, in retrospect, to be justified. Far from being musically irrelevant, the past decade of research has shown that low-level phenomena, such as the mechanics of the basilar membrane, have had far more impact on musical organization than was formerly suspected.
3. Thirdly, regarding the nay-saying character of much psychology of music research, history has largely vindicated the nay-sayers. For example, ongoing research on the perceptibility of serial transformations has been carried out since the 1950s. Careful, sophisticated experimental research has been carried out by scholars such as Bruner, Francès, Gibson, Lannoy, Largent, Millar, Pedersen, Thrall, and others. Yet, to

my knowledge, not a single one of these scholars has had his or her work cited by any set theorist. Many music theorists continue to write as though questions of perceptibility remain unaddressed and open. Some theorists wrongly assume that research has only addressed the listening of non-musicians or non-experts. (Gibson, for example, studied members of the Society for Music Theory.) Set theorists have been delinquent in ignoring this research. Music theorists in general have been delinquent when assuming that the human capacity for auditory experience is unbounded.

4. Finally, regarding the conservative musical tastes of music psychologists, it must be noted that the vast majority of music psychologists received their academic training in psychology, not in music. Music psychologists were no more conservative in their tastes than the general population. Many psychologists were notably supportive of new music (e.g. Francès). The more pertinent question is why more music scholars didn't make the effort to learn how to do psychological research. Fifty years ago, Farnsworth complained that few musicians were competent psychologists. That's just as true today as it was in 1948. If music psychology seems to favor a psychological perspective, that is largely because music scholars have generally failed to get involved. In fact, speaking now as a musicologist, I believe that musicology owes a collective debt of gratitude to the innumerable psychologists whose extraordinary efforts laid the groundwork for the discipline.

The Cognitive Revolution

Let's now turn to the second historical current contributing to cognitive musicology, the cognitive revolution.

The term "cognition" has many connotations. For the non-specialist, cognition is more or less synonymous with thought or thinking. Psychologists have used the term to designate various forms of knowing, and in some cases, psychologists have regarded cognition as equivalent to "the functioning of the mind."[2]

The rise of cognitive psychology is often traced to Ulric Neisser's book of that name, published in 1967. However, the origins of cognitive approaches to psychology can be seen in several earlier strands of research in psychology that led to increasing disgruntlement with behaviorism.

For most of the early part of the twentieth century, psychology, especially American psychology, was dominated by the behaviorist approach associated with J.B. Watson and (later) B.F. Skinner. Watson argued against positing mental states that were unnecessary for explaining a behavior. For example, the fact that an animal approaches a food dish does not mean that the animal has a *desire* or a *conscious intent* to eat. There is no way for an observer to "see" such a presumed conscious intent or desire.

To be fair, Watson's severe approach to psychological reasoning was a deliberate reaction against more informal psychological discourse whose theories appeared to be impossible to test. Watson and Skinner's behaviorism was simply an application of Occam's razor in the domain of mental processing. According to Skinner, we shouldn't posit sophisticated mental states when a simpler explanation can account for the experimental data equally well. This belief accounted for Watson's well-known (and notorious) disdain for appeals to consciousness as an unseen epiphenomenon, even in humans. Skinner, by contrast, never shared Watson's view regarding consciousness. Nevertheless, Watson and Skinner had much in common with the logical positivist, A.J. Ayer, and so it is not unreasonable to characterize behaviorism as "positivistic."

In our simplified story, the end of behaviorism's popularity can be loosely attributed to three events. First, experimental research itself implied the existence of higher-level mental processing that appeared to be essential in many tasks, especially those tasks that resembled natural problem-solving activities. Some psychologists, such as Broadbent, noted in their experiments that human subjects weren't simply *responding* to stimuli; they were anticipating and interpreting events, and different subjects appeared to be motivated by different goals. Increasing numbers of psychologists became interested in studying memory, attention, pattern recognition, concept formation, categorization, reasoning, and language. Behavioral methods seemed well suited to studies of sensation and perception, but behaviorism proved less useful in investigating more complex mental functions.

A second contributing factor was the advent of computer science and artificial intelligence. Computer programs were the very epitome of invisible information processors. In computers, the relationship between inputs and outputs depends critically on the nature of such invisible programs. Clearly, complex and multifaceted information processing functions can exist without anyone (apart from the programmer) knowing about their existence. If computer programs can be invisible yet real, then it is more plausible that analogous unseen mental functions can exist for humans and other animals.

Finally, a third influence was a general unhappiness with the reductionistic and simple mechanistic view of mental life that was implied by Skinner's work.

In contrast to behaviorism, the new cognitive psychology could be characterized by three dispositions. First, there was a willingness among cognitive psychologists to entertain explanations of mental processes and mental states that could not be behaviorally observed. In effect, some intellectual space was made for plausible invisible mental functions -- the sort of functions that might provide motivations, such as initiating actions, rather than simply responding to a stimulus. Second, there was a consensus that a useful way to study the operation of the mind is to decipher and describe underlying mental representations. That is, cognitive psychologists became interested in how skills, perceptions, knowledge, beliefs and motivations might be mentally coded, stored and retrieved. Third, cognitive psychologists placed special emphasis on the *processes* of thought instead of its *content.* [3]

In the early years, cognitive psychology tended to eschew psychophysics, sensation, and neural aspects of mental behavior. However, in recent decades, cognitive psychologists have shown a renewed interest in the mechanisms of mental life. Where formerly cognitive psychologists were interested in discussing mental life and mental functions apart from mechanisms, in recent years, cognitive psychology has connected once again to those perceptual and biopsychology researchers who remained tied to behaviorist methods. This integrative tendency is reflected, for example, in the burgeoning field of cognitive neuroscience.

In retrospect, cognitive psychology has prevailed over behaviorism, primarily because behaviorism fell prey to what is now referred to as the *positivist fallacy.* If a phenomenon results in no observable behavior, a researcher may be tempted to wrongly conclude that no mental activity has taken place. In short, the positivist fallacy arises when *absence of evidence* is mistaken for *evidence of absence.* We will return to the issue of the positivist fallacy again in my third lecture on methodology where we will see that this fallacy has plagued not only scientific research, but humanities scholarship as well.

What is Cognitive Musicology?

At this juncture, we might offer a preliminary definition of cognitive musicology. Cognitive musicology is an area of musicology that studies musical "habits of mind." It is a field that has been inspired by the cognitive revolution and informed by past lessons and mistakes in the psychology of music. In contrast to the behaviorists, cognitive musicologists do not presume that there is a simple relationship between stimulus and response. Musical stimuli and the phenomenal experiences they evoke typically have sophisticated, complex, and mostly unobserved mental functions interposed between them. Cognitive musicologists are primarily interested in processes rather than content. We accept that listeners, performers, composers, improvisers, dancers and others have specialized knowledge, beliefs, motivations, skills and strategies. We tend to focus on mental representations for music, but we don't regard these representations as disembodied abstractions: musically pertinent representations are concretely expressed in human biology and often exist as socially distributed codes as well. In investigating the musical mind, it is not the task of the cognitive musicologist simply to document limitations to musical experience, but also to point to the unexplored cognitive terrain -- regions of musical possibilities that have not yet been visited by creative artists.

In summary, music cognition is an approach to the study of music that places the mind in the central position. To study music is to study the musical mind.

Mental Representations of Music

As I've just noted, a major preoccupation for cognitive musicologists is the study of mental representations for music. Music-lovers will have no difficulty believing that most of what is musically valuable is unobservable -- at least not observable with the unaided or untutored eye. Experienced performers, for example, know all too well that there is hardly any difference in facial expression between those members of an audience who are in rapture, and those who would rather be somewhere else. However, the presumption that cognitive processes are difficult to observe is open to abuse. As the behaviorists rightly fear, one might claim that all sorts of spurious processes exist. Whenever possible, the cognitive musicologist needs to demonstrate that a presumed music-related mental representations does, in fact, exist. Let me illustrate some mental representations by invoking some specific examples.

EXAMPLE 1: Musical Memory

As quickly as you can, I want you to answer the following question, *yes* or *no:*

> Does the word "but" occur in the lyrics to the song *Row, Row, Row Your Boat?*

[This example doesn't work if the reader doesn't actually try the task.]

If you are familiar with the song, you probably solved this problem by scanning the lyrics from the beginning of the song. More precisely, you probably mentally generated a speedy rendition of the work until you encountered the word "but" in the phrase "life is but a dream" and then you stopped searching. There are at least three conclusions we can draw from this little task:

1. We are able to access mental representations for music. In this case, I had you focus on the lyrics, but the same can be done for melody alone.
2. We can access music-related representations in the total absence of sound.
3. We can manipulate these mental representations in certain ways (such as speeding up the rendition beyond what would be musically acceptable). But we cannot manipulate these mental representations in any way we wish. For example, you might have been able to answer my question much more quickly if you had random access to all of the words of the lyrics. Similarly, it would have been faster if you could start at the end of the lyrics and work your way forward. Either of these two strategies would have generated a faster answer to my question, but as far as we know, people are unable to do this. It is as though the mental representation for *Row, Row, Row Your Boat* is a linear recording that we must play from the beginning (or from a handful of possible starting points). Once again, my third point here is that we can access and manipulate musical representations only in certain ways.

EXAMPLE 2: Perceptual Schemas

Let's consider now a second example that requires a little more musical sophistication. Sing any tone to yourself. Now I'd like you to hear this pitch as a tonic pitch (or `doh') in a scale that begins on that pitch. In fact, if you are like most people, you already would have been hearing this pitch as a tonic even before imagining the scale.

Let's now have you hear this same pitch differently. Once again, sing the pitch, only this time I want you to hear this pitch as the dominant scale degree (or `so'). Now, for those of you who are able, try hearing the same pitch as the leading-tone (`ti'). Now hear it as the mediant pitch (`mi'). Notice how much longer it takes to hear the pitch as `mi' compared with `doh'.

Figure 1 shows response-time data for five music students. Each musician heard a randomly selected tone, and was asked by a computer to hear the tone as a particular scale degree. We then measured how long it took our listeners before they responded that they were hearing the tone in the specified way. In order to be certain they weren't fibbing, we then played a cadence and asked them to indicate whether or not the cadence corresponded

with the imagined key. The data in Figure 1 plot the results only for correct responses.

Figure 1

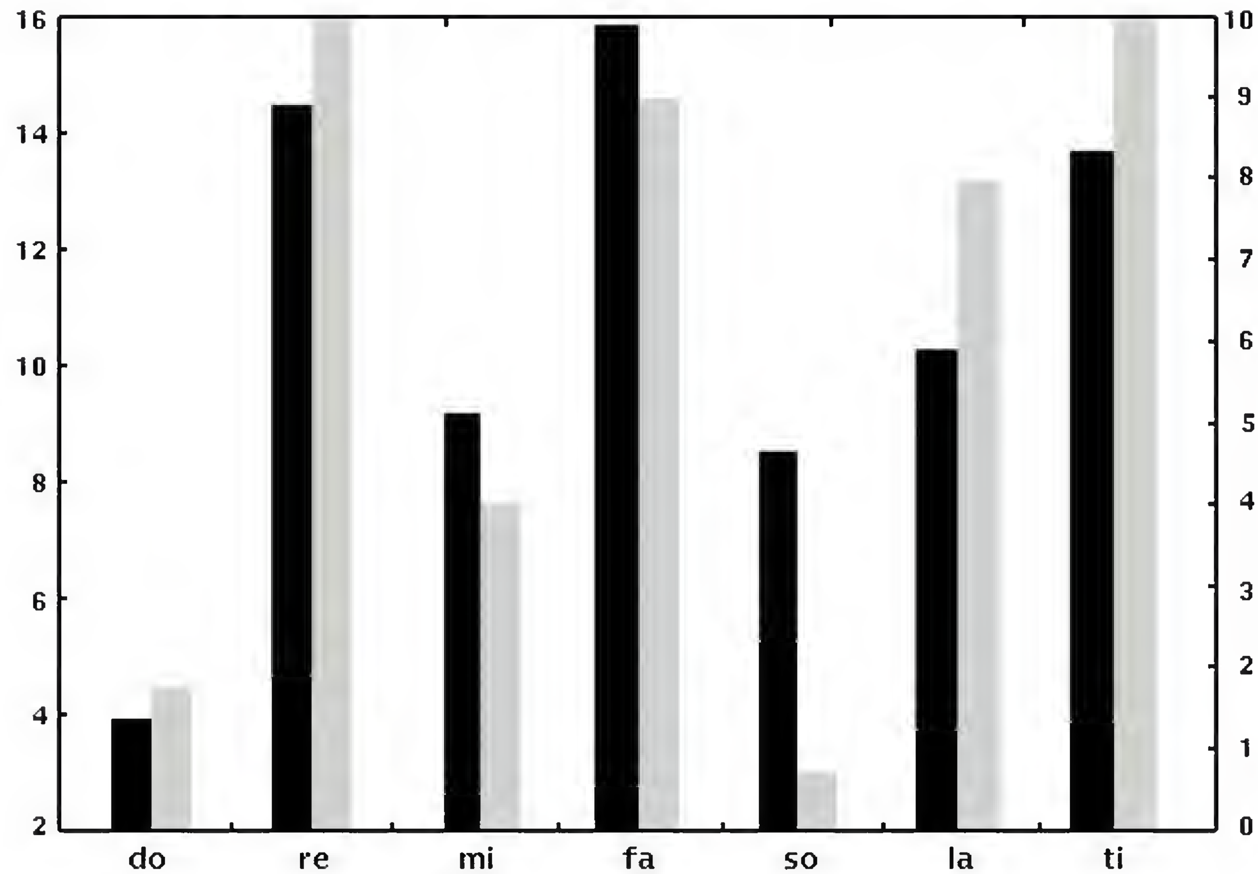


Fig. 1: Median response times for scale degree orientations. Black bars indicate the median response time for imagining a tone as the specified scale degree (left scale in seconds). Grey bars indicate the frequency of occurrence for various traditional folksongs beginning with the specified scale degree (right scale in bits).

You can see that hearing a tone as the tonic takes the least amount of time. Hearing the tone as the dominant is the next fastest. Perhaps surprisingly, hearing the tone as a subdominant (`fah') takes the longest time to imagine.

We know from other research in psychology that *response times* (how long it takes to do something) tell us something about how much mental effort is involved in the task. (A classic illustration of this is Roger Shepard's famous work on mental rotation.)

Response times tell us something about the complexity of the mental representation. For an isolated pitch, the least mental effort is required to hear that note as a tonic. In fact, we know that people who don't have perfect pitch unconsciously presume that an isolated pitch is a tonic. It requires considerably more effort to hear that note as a non-scale tone.

There are again, several conclusions we can draw from this brief illustration:

1. There is a difference between *hearing* and *hearing as*. Any person with normal hearing can hear a tone, but not everyone can hear the tone as (say) "fah."
2. *Hearing as* is a natural tendency when hearing tones. The existing cognitive research suggests that listeners automatically and unconsciously make assumptions about the scale context (or what musicians call "tonal function") of a pitch.

3. Some *hearing as*'s are easier to hear than others. For example, it is easier to hear an isolated tone as a tonic than to hear it as a mediant pitch. Once again, these tendencies reflect different aspects of mental representations. Reaction time provides a useful indication of the complexity of mental processing.
4. *Hearing as* is obviously related to one's cultural background. The vocabulary of scale degrees is passively learned from the cultural milieu. For most of the people in this room, it is simply impossible to hear a tone as the pitch *hwang* in a traditional Korean scale. Most of us haven't been exposed to the pertinent music.
5. Although I haven't presented any detailed evidence, another conclusion we can offer is that listeners are different. Of course people in different cultures are exposed to different musics -- and so they differ. But even within a single culture, differences of exposure are evident. An obvious example occurs for *absolute* or *perfect pitch*. Some people will be able to represent a sound by an absolute pitch name (e.g., G#). But there are many other more subtle differences as well. The experimental evidence shows that not everyone listens in the same way, or has the same phenomenal experience.

EXAMPLE 3: Rhetorical Listening

Let's consider now an even more sophisticated example of a music-related mental representation: in this case, another form of *hearing as*. From the early Middle Ages until recent times, it has been common for musical commentators to relate music to rhetoric. Theorists like Heinrich Koch have suggested that musical materials can manifest different "tones of voice" or rhetorical character. In particular, Koch noted that the different formal sections in musical works can be characterized by such rhetorical differences. Using contemporary terminology, we can distinguish types of passages such as the following:

> **Closing material.** A closing passage conveys a feeling of impending finality. Such passages suggest that the work is ending, or that the end of the work may be expected shortly.
>
> **Expository material.** Expository passages present the basic musical ideas of a work, such as the principal melodies or themes.
>
> **Developmental material.** Developmental passages convey musical ideas that have been varied, broken up, or rearranged in some manner.
>
> **Transitional material.** Transitional passages act as links or bridges between other passages. They provide an interlude or prepare for something new.

We might well ask whether listeners are capable of hearing passages according to these rhetorical categories. To this end, Mei Yen Ch'ng, Kim Rasmussen and Sarah Stockwell and I recruited forty-three listeners. We assembled a number of brief passages (lasting 20 seconds each) taken from recordings of string quartets by Haydn and Mozart. The sample passages were randomly selected from sections that had already been analytically identified as the introduction, exposition, or development in a sonata-allegro movement. Transitional passages were randomly extracted from appropriate points in the exposition.

The listeners fell into three groups: music majors who had taken a course whose curriculum stressed the identification of music-rhetorical devices in symphonic works, a second group of matched music majors who hadn't taken such a course, and a third group of non-musician university students who claimed to have little or no formal musical background.

We found that listeners were able to identify all rhetorical categories significantly better than chance. As you might expect, "closing" passages were most easily identified, even though these passages never included a final chord or cadence. "Transitional" passages proved to be the most difficult to identify. We were surprised to find that all three groups of listeners were equally adept; the musicians were not better than the non-musicians. In fact, the raw scores for the non-musicians were slightly better than for the musicians, mostly because musicians showed a slight reluctance to classify passages as "transitional."

What does this mean? First, it suggests that listeners are indeed broadly capable of hearing brief musical excerpts in terms of rhetorical categories traditionally distinguished by music scholars. These rhetorical categories are psychologically salient; they make sense to people, they aren't merely formal abstract concepts. Moreover, this way of listening appears to be equally accessible to musicians and non-musicians. In the course of our experiment, we were pleasantly struck by how unphased our non-musicians were. They didn't receive any feedback, and we didn't give them any practice trials. Without ever having taken a music course, they seemed perfectly happy to classify passages as transitional, or developmental, or whatever. Most importantly, note that the test passages were presented in isolation, entirely removed from their musical contexts. There is something about (say) a development passage that sounds "developmental" even when the rest of the piece is unknown. Finally, since none of the passages used in this experiment straddled boundaries between formal sections, the results also imply that it isn't necessary to recognize sectional boundaries in order to follow the formal outline for a sonata-allegro work.

Cognition and Conscious Thought

We have just looked at three examples illustrating mental representations for music, namely memory for musical lyrics, perceptual schemas for hearing scale degrees, and hearing musical passages in terms of rhetorical categories.

It isn't often we get asked whether the word "but" occurs in the lyrics of some song, or to hear a particular pitch as some specified scale degree. It would be useful to know, not just what people are *capable* of doing, but also what they commonly or typically do. In particular, since the word "cognition" implies some sort of "cogitation" or conscious "thinking" we might ask *what do people typically think about when they listen to music?* Unfortunately, this isn't easy to answer.

In 1994, I made a preliminary effort to try to answer this question. I was teaching two sections of the same course in music theory. Each class consisted of roughly 30 students. In the first class I distributed a questionnaire which remained face down on their desks while they listened to two minutes of music. The music was a segment from a Mozart symphony, selected at random. After the music ended, the students turned over their questionnaires. The questionnaire began as follows:

> "You have just listened to two minutes of music. The purpose of this questionnaire is to have you report on what you were thinking about during this time. Please answer the questions honestly. The questionnaire is intended to be anonymous, so do not write your name on this paper."

Students were asked a series of questions; they were asked to estimate the proportion of time they spent on certain types of activities. The most commonly reported activity was *thinking about things I have to do today.* Students were encouraged to provide written elaborations on the reverse side of the questionnaire.

I repeated this same informal experiment with the second section of the same music course. This time, I played the same recording, but with the amplifier turned off. That is to say, the entire class sat in silence for two minutes. (Incidentally, that's a long time for a group of people to sit in silence.) After the two minutes had elapsed, this second group of students were similarly asked to answer a questionnaire.

> "You have just sat in silence for two minutes. The purpose of this questionnaire is to have you report on what you were thinking about during this time. ..."

As you might expect, these students reported a wealth of daydreaming scenarios.

I then compared the responses of the two groups of students. As expected, the group that listened to the Mozart symphonic passage reported significantly more music-related thoughts. But the size of this difference was tiny. On average, the group exposed to the music reported less than 5 percent of their thoughts related to music, while the non-exposure group reported only 1 percent of their thoughts related to music. This means that, over the 120

seconds of music, the group that listened to the music spent on average about 6 seconds thinking about the music. In effect, the typical student's thinking went something like this:

> "This sounds like Mozart, maybe Haydn but probably Mozart. A symphonic work, no solo instrument so not a concerto. Um, what should I do after school tonight? ..."

Six seconds of music-related thought, and then they were gone for the next 114 seconds. And this occurred in a music theory class, where a music professor had handed out a questionnaire that could well have been a surprise quiz.

There are a number of methodological problems with experiments such as this that rely on introspection, especially when we try to assess unguided mental activity. But this informal experiment is nevertheless suggestive. It implies that the predominant conscious mental activity engaged in while listening to music is daydreaming.

Since research has established that listening to music entails a host of mental representations (see, for example, Krumhansl, 1990), the corollary of listener-daydreaming is that most music-related mental representations must be unconscious phenomena. Although most people in industrialized countries are exposed to lots of music, it appears that they don't think many music-related thoughts while listening.

Listening Strategies

Of course not *all* listening is unconscious or pre-verbal. Listeners may approach a listening experience with different strategies or different mental habits at different times. Elsewhere I have written about listening styles and listening strategies and have described some 20-odd common approaches to music listening. Let me give you the flavor of these by describing just one listening style, which I call *fault listening*. It is a listening mode that has a strong conscious component.

For several years I lived in the United Kingdom, and while there I was a perennial listener to the BBC's classical music network known as Radio 3. Unlike radio broadcasting in North America, European classical programming relies much less on commercial recordings. At the time that I lived in Britain, the majority of classical radio programming entailed live or delayed-live broadcasts.

As a listener accustomed to hearing virtually flawless commercial recordings, I vividly recall the shock of hearing performers make mistakes on the radio. What I found remarkable was how the occurrence of a single mistake would utterly transform my listening. Having heard one mistake, I was "all ears" -- vigilant to identify further errors or lapses of musical judgment.

Fault listening might be defined as follows: it is a listening mode that arises when the listener is mentally keeping a ledger of faults or problems. A high-fidelity buff may note problems in sound reproduction. A conservatory teacher may note mistakes in execution, problems of intonation, ensemble balance, phrasing, etc. A composer is apt to identify what might be considered lapses of skill or instances of poor musical judgment.

Fault listening tends to be adopted as a strategy under three circumstances: (1) where an obvious fault has occurred, the listener switches from a previous (often passive) listening mode and becomes vigilant for the occurrence of more faults; (2) where the role of the listener is necessarily critical, as in teachers, conductors, or music critics; or (3) where the listener has some prior reason to mistrust the skill or integrity of the composer, performer, conductor, audio system, etc.

There are many other listening styles and strategies we could discuss, but we don't have time. This single example should suffice to establish my point. Even as individual listeners, we have a palette of different ways to approach the listening experience. In some cases we can switch strategies in the middle of a musical work. As individuals, we undoubtedly have preferred ways of listening; some arise from enculturated habits, some from professional training, and others from personal disposition or mental habit.

Investigating Musical Thought

Let's pause for a moment and take stock. As we have seen, cognitive musicology is predominantly the study of musical thought and mental representations. We've seen three examples in memory for musical lyrics, schemas for hearing scale degrees, and hearing musical passages in terms of rhetorical categories. We've also encountered evidence suggesting that most music-related mental phenomena are unconscious in nature. But we've also seen an example of a more conscious listening style in strategies such as "fault listening."

All of these examples have related to listening, and all have relied on introspective accounts of our mental experiences. In the time remaining, I'd like to broaden our discussion and address five more extended examples that are intended to highlight several contrasts. The examples include both socio-cultural phenomena and neurological phenomena; they address historical, performance, compositional, and listening issues; the repertories span archaic to contemporary popular music, and include cultures from five continents.

1. Musical Notation: Deciphering an Ugaritic Song

How do we gain access to the minds of people and cultures long past? We have no direct access to their thoughts, but that's also true of people sitting right next to us. We can glimpse mental activities by examining whatever externalized evidence is available. In some cases, the available evidence can be very small. Consider the oldest known musical notation, shown in Figure 2.

Figure 2: Ugarit music tablet.

In 1929, the French archaeologist Claude-Frédérric-Armand Schaeffer began a series of excavations at Ras Shamra on the Mediterranean coast of Syria. Schaeffer uncovered hundreds of clay tablets bearing testimony to the ancient city of Ugarit, a site that was home to a succession of cultures from the 6th to the 1st millennium BC. The document reproduced in Figure 2 comes from the most prosperous age in Ugarit's history and is dated between 1450 BC and 1200 BC.

The text uses cuneiform writing organized from left to right. The language is Hurrian, a language that has largely been deciphered. However, this particular tablet (and several others like it) have so far resisted complete decipherment. Laroche (19XX) observed [pp. 462f., 484] that the section above the double line forms a coherent text that contains several repetitions resembling refrains found in musical lyrics or poetry. Below the double line is a combination of words and numbers. Hans Güterbock (1970) noted that the words are Hurrian equivalents to [Sumerian??] musical terms that had already been deciphered. Specifically, the terms indicate the names of the intervals formed by strings of a 9-stringed harp or lyre. In the Ugarit tablet, each interval term is followed by a single number (refer to Figure 3).

Figure 3: Ugarit transcription (text).

There are at least six modern attempts to transcribe this work into contemporary Western notation. The most difficult challenge has been interpreting the meaning of the numbers following each interval term. Do these numbers represent the number of repetitions of the intervals, or the number of upward scale tones from the lower to the upper string of the interval, or the number of downward scale tones from the upper to the lower string of the interval?

Figure 4: Two interpretations of the Ugarit tablet.

Figure 4 shows excerpts from two different transcriptions, one by XXX and the other by Anne Draffkorn Kilmer. It's hard to imagine more contrasting decipherments.

Now I'm not at all an expert in Ugarit, nor am I a historical musicologist. However, what we know from music cognition may be of some help in deciphering the music. Consider, for example, the finding by Vos and Troost (1989) that showed that most large intervals in melodies ascend in pitch. That is, intervals such as perfect fifths and major sixths are significantly more likely to rise than fall.

Figure 5 illustrates this phenomenon for a number of repertoires I've examined, including songs from the following cultures: Arabic, Austrian, Belgian, Czech, Dutch, English, French, German, Italian, Yugoslavian, Russian, Spanish, Chinese, Korean, Japanese, Hassidic, Ojibway, Tahitian, Pondo, Venda, Xhosa, and Zulu. In addition, I've examined American popular songs, Schubert Lieder, and Gregorian chant. In all of these repertories, there is a significant tendency for large pitch intervals to ascend rather than descend. We don't yet know the reason for this phenomenon; however, it might be related to pitch "declination" in speech.

Figure 5

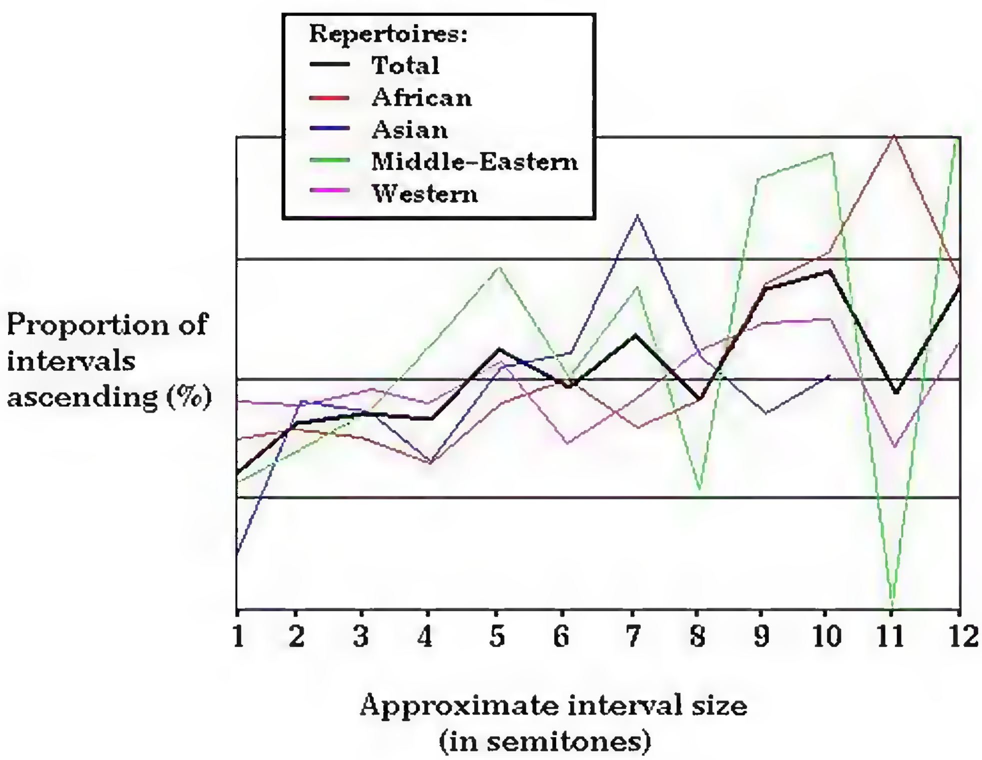


Fig. 5: Proportion of ascending/descending intervals for 22 cultures. In general, most small intervals tend to descending whereas most large intervals tend to ascend. Cultures included: Arabic, Austrian, Belgian, Chinese, Czech, Dutch, English, French, German, Gregorian chant, Italian, Korean, Japanese, Hassidic, Pondo, Russian, Spanish, Tahitian, Venda, Xhosa, Yugoslavian, Zulu. N.B. In all cultures, intervals roughly 11 semitones in size tend to be rare, hence the corresponding plotted values have a low reliability.

Such a pattern in no way proves that large leaps are more likely to ascend than descend in Ugarit music. But a predominance of descending large leaps would certainly be unusual given our knowledge of other musical cultures.

Unfortunately, time doesn't permit a complete enumeration of the discoveries about melodic organization that might be pertinent to deciphering the Ugaritic tablets. Suffice it to say that there are at least a dozen features of melodic organization that have been established through systematic study, and these principles could provide independent evidence in support of some proposed transcriptions at the expense of others.[4]

2. Transcultural and Historical Listening: The Case of Melodic Accent

A question that has long preoccupied ethnomusicologists is the extent to which we can hear the music of another culture in the same manner as culturally-experienced listeners. In fact, this question is a central issue in historical musicology as well. Even if we were to hear period-authentic sound recordings, we might well ask whether the modern listener experiences the music in a manner similar to past listeners.

In order to consider this question, we need to distinguish many possible aspects of musical experience. A modern listener might hear the pitches the same way as a past listener, but not hear the connotations of the timbres in the same way. A modern listener might apprehend the musical program or context, yet fail to hear the radical betrayals of harmonic expectations. In other words, we need to ask to what extent a modern listener can have an experience similar to a past listener for each of several aspects of musical experience.

For illustrative purposes, let's focus on one aspect of musical behaviors, whether modern and past listeners experience accent (or stress) in a similar way. Over the centuries, music theorists have proposed a number of factors which are thought to contribute to stress or accent in music. For example, accents are presumed to arise through increased loudness ("dynamic accent") and through increased duration ("agogic accent"). One of the most contentious forms of accent has been the notion of pitch-related accent, or *melodic accent.* Some theorists have suggested that higher pitches are more accented than lower pitches (you'll find this view, for example, in Benward and White). Other theorists (such as Parncutt) have argued the reverse: that low pitches are perceived as more accented. Yet other theorists proposed that both extremes of high and low pitch are more salient than mid-register pitches. Other theorists, for example Graybill, have claimed that it is the size of the interval that's important: large intervals are more accented than small intervals. Some (such as Rothgeb) have suggested that it is only ascending intervals that are important. Other theorists, notably Joel Lester, have argued that it is not pitch height or interval size that's important, but rather changes of melodic contour -- that is, pivot points in a melody.

For modern listeners these different notions of melodic accent have been tested experimentally by Woodrow and by Squire. Unfortunately, the perceptual evidence indicates that modern listeners do not experience any of these forms of presumed melodic accent. Of course it is possible, that listeners in different historical periods heard melodic accent differently. Without any knowledge of these modern experiments, the theorist William Caplin was surprisingly prescient when some years ago he questioned whether any of these ideas of melodic accent hold merit. In 1982, the Dutch researcher Joseph Thomassen carried out two sets of perceptual experiments and formulated what is now regarded as the best model of melodic accent (for modern listeners); unfortunately, it's a model that's too complicated to describe succinctly, so I'll skip the details here.

In 1996, Matthew Royal and I published the results of a series of studies testing eight different notions of melodic accent. Instead of approaching the problem by carrying out further perceptual experiments, we decided to study a large sample of notated music to measure which concept was most consistent with how composers actually compose. We studied three contrasting repertoires of music containing a total of two hundred works. Although the works spanned a considerable historical period, in all three repertoires, we found that Thomassen's model was significantly superior to all the other proposed notions of melodic accent of which we are aware.

What's important from a historical point of view is that one of the repertoires we tested was a sample of Gregorian chant. Now in most music, different types of accent tend to coincide -- accent types tend to be synchronized. That is, notes which have longer durations tend to be given greater dynamic accents, and both of these tend to occur in stronger metric positions. In addition, when the music has some sort of text or lyrics, the accents tend to coincide with syllable onsets rather than with a sustained syllable, or mellisma. This tendency to synchronize accent types is illustrated in Figure 6 where agogic (duration), metric, dynamic, melodic (contour), and syllable onset are all coordinated.

Figure 6

Figure 6: Synchrony of accent types. Agogic (duration) accent, metric accent, dynamic accent, melodic (contour) accent, and syllable onset are all coordinated.

Matthew Royal and I found that the tendency for accent types to be synchronized also holds true for melodic accent. By and large, melodic accents tend to occur in strong metric positions, are associated with longer duration notes, receive more dynamic stress, and tend to coincide with syllable onsets rather than with sustained syllables. The exceptions to this generalization occur for syncopated and hemiola passages where one or two accent types are systematically offset from the others.

Royal and I were surprised to discover a notable exception in the case of Gregorian chant. As in the other repertoires, in the chant literature, there are marked correlations between the occurrence of melodic accents (as defined by Thomassen's model) and whether or not the moment is syllabic or mellismatic. However, the correlations are *negative* rather than positive. Pitches that are deemed to convey a melodic accent are much more likely to occur on a mellisma than at a syllable onset. Let me try to illustrate this using *Happy Birthday.* In Figure 7, I've miscoordinated the syllable placement with respect to metric position and agogic accent:

Figure 7.

Fig. 7: *Happy Birthday* re-texted in order to reduce the correlation between syllable onsets and strong metric positions.

In chant, the miscoordination is between syllable placement and melodic accent. The miscoordination is utterly systematic. Of the 60 randomly selected chants we studied, only a single chant did not display this methodical miscoordinated relationship between melodic accent and text. In the first instance, this suggested that the musicians who created or subsequently modified these works were purposely trying to avoid highly stressed or inflected moments in the music.

Some musicologists (a small minority) have suggested that chant might have been originally sung in a rhythmic fashion (and that modern arrhythmic performance of chant is somehow an aberration). However, the statistical correlations do not at all support this view.

Incidentally, the single exception in our sample of chant was *A Solis Ortus Cardine,* the text of which is given in Figure 8. The syllable stresses as published in the *Liber Usualis* are also shown, as well as a simple representation of the stress pattern. One can clearly hear the iambic tetrameter rhythm here; the poetic text is highly rhythmic:

Figure 8: Text for *A Solis* [Liber Usualis, p. 400; #12].

A solis ortus cardine	A so/- lis or/- tus car/- di- -ne	. > . > . > . .
ad usque terrae limitem,	ad us- que ter/- rae li/- mi- tem,	. > . > . > . .
Christum canamus principem,	Chri/- stum ca- na/- mus prin/- ci- pem,	> . . > . > . .
natum Maria Virgine.	na/- tum Ma- ri/- a Vir/- gi- ne.	> . . > . > . .
Beatus auctor saeculi	Be- a/- tus au/- ctor sae/- cu- li	. > . > . > . .

servile corpus induit:	ser vi/- le cor/- pus in/- du- it:	. > . > . > . .
ut carne carnem liberans,	ut car/- ne car/- nem li/- be- rans,	. > . > . > . .
ne perderet quos condidit.	ne per/- de- r et quos con/- di- dit.	. > . . . > . .
Castae parentis viscera	Ca/- stae pa- r en/- tis vis/- ce- ra	> . . > . > . .
cae lestis intratgratia:	cae/ le/- stis in/- trat- gra/- ti- a:	> > . > . > . .
venter puellae bajulat	ven/- ter pu- el/- lae ba/- ju- lat	> . . > . > . .
secreta, quae non noverat.	se- cre/- ta, quae non no/- ve- rat.	. > . . . > . .
Domus pudici pectoris	Do/- mus pu- di- ci pe/- cto- ris	> . . . . > . .
tem plum repente fit Dei:	tem/ plum re- pen/- te fit De/- i:	> . . > . . > .
intacta nesciens virum,	in- ta/- cta ne/- sci- ens vi/- rum,	. > . > . . > .
concepit alvo filium.	con- ce/- pit al/- vo fi/- li- um.	. > . > . > . .

Now I'm not a chant scholar, so I know nothing about the origin of this work. But even if we didn't know that the *text* is rhythmic, the synchronization between the syllable placement and what we know of perceived melodic accent (for modern listeners) suggests that it is indeed likely that this particular work was sung rhythmically and that it differs significantly from the other chants we studied.

When Royal and I did this work, we were also struck by something else. Joseph Thomassen's model of melodic accent was formulated from tests using Dutch listeners in the early 1980s. In carrying out our statistical analyses we found that the relationship was significant at less than one chance in a million. That is, there is less than one chance in a million that a handful of modern Dutch listeners sitting in a laboratory listening to sequences of sine tones would respond in a way that corresponds to the text setting of music created roughly a thousand years ago. Moreover, this robust correlation was found only for Thomassen's model of melodic accent. Other conventional views of accent (such as the highest pitches, the largest intervals, etc.) did not show such correlations -- and let me remind you that the existing perceptual research is consistent *only* with Thomassen's model.

The inescapeable conclusion is that, whatever melodic accent is, it doesn't seem to have changed much over the past millennium. Modern listeners may not hear Gregorian chant the same way that Medieval listeners do, but we appear to hear the melodic accents in a similar way.

Where historical musicologists might infer rhythmic performance based on source studies, rescension, and other standard techniques, it seems that cognitive musicology might well be able to provide independent corroborating evidence of a particular interpretation of the music of the past. The research also might assist scholars in distinguishing sub-repertoires that are often mixed together in the sources we have available for study. As Katherine Bergeron has shown, collections of such works can have unusual and sometimes bizarre origins.

3. Performance and Idiomaticism

A common mistake is to regard cognitive representations of music as arising solely from the perception of music. However, there are many cognitive aspects of music that have nothing to do with perception. Good examples of non-perceptual phenomena that are reflected in musical organization can be found in performance idiomaticism. Since music is often performed using musical instruments, the mechanics of the instruments themselves often influence how the music is structured.

Some of these performance aspects are relatively easy to identify. A trivial example occurs when a musical work is composed to lie within the pitch range of some particular instrument. Another obvious example is evident in the contrast between wind instruments and non-wind instruments. When composing for French horn, for example, the composer must accommodate the performer's need to breathe by providing periodic rests. A work composed for 'cello is often impossible to perform on (say) the bassoon, because the bassoonist is constantly trying to find a place to breathe.

Other idiomatic aspects of performance are less directly observable, though still evident. Ethnomusicologists have frequently observed that instrumental idioms appear to have marked impacts on the character of music-making in different cultures (e.g., Yung, 1980; Baily, 1985; Kippen & Bell, 1989). Similarly, jazz musicians have often stressed the importance of idiomatic instrumental techniques in improvisation (e.g., Sudnow, 1978, 1979).

The most distinctive instrumental idioms are those gestures that are unique to a given instrument. For example, a well-known solo trumpet passage at the end of Leroy Anderson's *Sleigh Ride* imitates the sound of a neighing horse. This effect is almost impossible for any other instrument to produce, and so the relative ease with which it can be done on the trumpet means that it is justifiable to characterize the gesture as "idiomatic to the trumpet."

More subtle instrumental idioms are evident in a study of works for trumpet carried out by myself and Jonathon Berec in 1993. Berec and I began by collecting detailed performance data from two performers, one professional and one amateur. The measurements included many of the mechanical aspects of performance, including fingering, tonguing, embouchure, and breathing techniques. For example, the trumpet performers were asked to tongue notes as rapidly as possible in different registers and at different dynamic levels. Measurements were taken of how long the performers could sustain tones, and how quickly they could inhale. In addition, measurements were made of the speed of loss of muscle tone in the embouchure for sustained playing. Data was also collected on the difficulty associated with pitch movements within registers. In the case of fingering difficulty, the trumpet players themselves estimated the degree of difficulty for all possible transitions between two successive finger/valve combinations. The following table shows the average degree of difficulty for each of the possible finger/valve transitions, as judged by our two performers. Rows and columns represent antecedent and consequent finger/valve positions. For example, on a scale of difficulty ranging from zero to ten, the transition from first valve (1) to second and third valve (2-3) received an average rating of 7.5.

Table 1.

Mean difficulty for finger/valve transitions as judged by two trumpet players.

Valve combination for the consequent tone.

	0	1	2	3	1-2	1-3	2-3	1-2-3
0:	0.0	1.0	1.0	1.9	1.5	3.0	3.0	3.5
1:	1.0	0.0	2.0	3.0	2.0	4.5	7.5	6.0
2:	1.0	1.5	0.0	5.3	3.0	9.5	6.0	9.0
3:	2.5	4.0	4.5	0.0	7.0	4.0	4.0	5.5
1-2:	1.5	1.5	2.3	7.5	0.0	6.0	6.0	5.0
1-3:	3.5	4.0	9.5	1.5	5.5	0.0	6.0	4.0
2-3:	2.5	6.0	5.5	4.0	5.0	5.5	0.0	3.8
1-2-3:	3.0	4.0	8.5	3.5	6.0	5.0	5.0	0.0

Having collected all of this data, we constructed a computer model of the trumpet/performer interaction. For any given musical score or passage, the model is able to generate estimates of the degree of difficulty for each of seven technical aspects of performance: (1) pitch register, (2) dynamic level, (3) fingering, (4) tonguing, (5) embouchure endurance, (6) breathing, and (7) intervallic transitions. We tested the model by comparing the difficulty estimates with graded trumpet études from a well-established conservatory curriculum.

After developing our trumpet model, we applied it to several trumpet works. Some works were written by trumpet virtuosi while other works were written by non-trumpet players. The virtuoso works included Malcolm Arnold's *Fantasy* for trumpet, Guillaume Balay's *Prélude et ballade*, and Herbert Clarke's *Stars in a Velvety Sky*. In addition, the three movements of Paul Hindemith's trumpet sonate were examined.

Idiomaticism

Just because a work is easy to perform on a given instrument does not make it idiomatic to that instrument. The work may be easy to perform on all instruments. A gesture is idiomatic when it is can be produced with *comparative* or *relative* ease. That is, given what *could* be the case, the actual arrangement renders the music more manageable.

Consider, by way of example, the effect of key on performance difficulty. Suppose we were to transpose a work through all twelve pitch-classes, and compare the difficulty for all keys. If a work was written in the key of Eb major, and Eb major turned out to be the most difficult of all possible keys, then we could not claim that the work is idiomatic to the instrument. On the other hand, if we found that the key of Eb major exhibited the lowest difficulty score, then this would lend weight to the claim that the work was created with the instrument in mind.

The following two graphs show the effect of transposition on fingering difficulty estimates for the Arnold, Balay and Clarke works. Notice, first of all, that the fingering difficulty shows a general tendency to fall as the work is transposed up in pitch. Brass players will recognize that this is a simple consequence of the way the harmonics and valves interact. As a work is transposed higher, there is less need to use some of the more difficult finger combinations.

Superimposed on this general downward trend you can see local fluctuations in difficulty depending on the key. The point marked zero along the horizontal axis represents the original key in which each work was written. You can clearly see that, with one exception, there is a notable minimum present. (The one exception is the slow second movement in Arnold's trumpet concerto.) The predominance of local dips at zero transposition suggests that the composers chose a key that facilitates performing the work.

Figure 9.

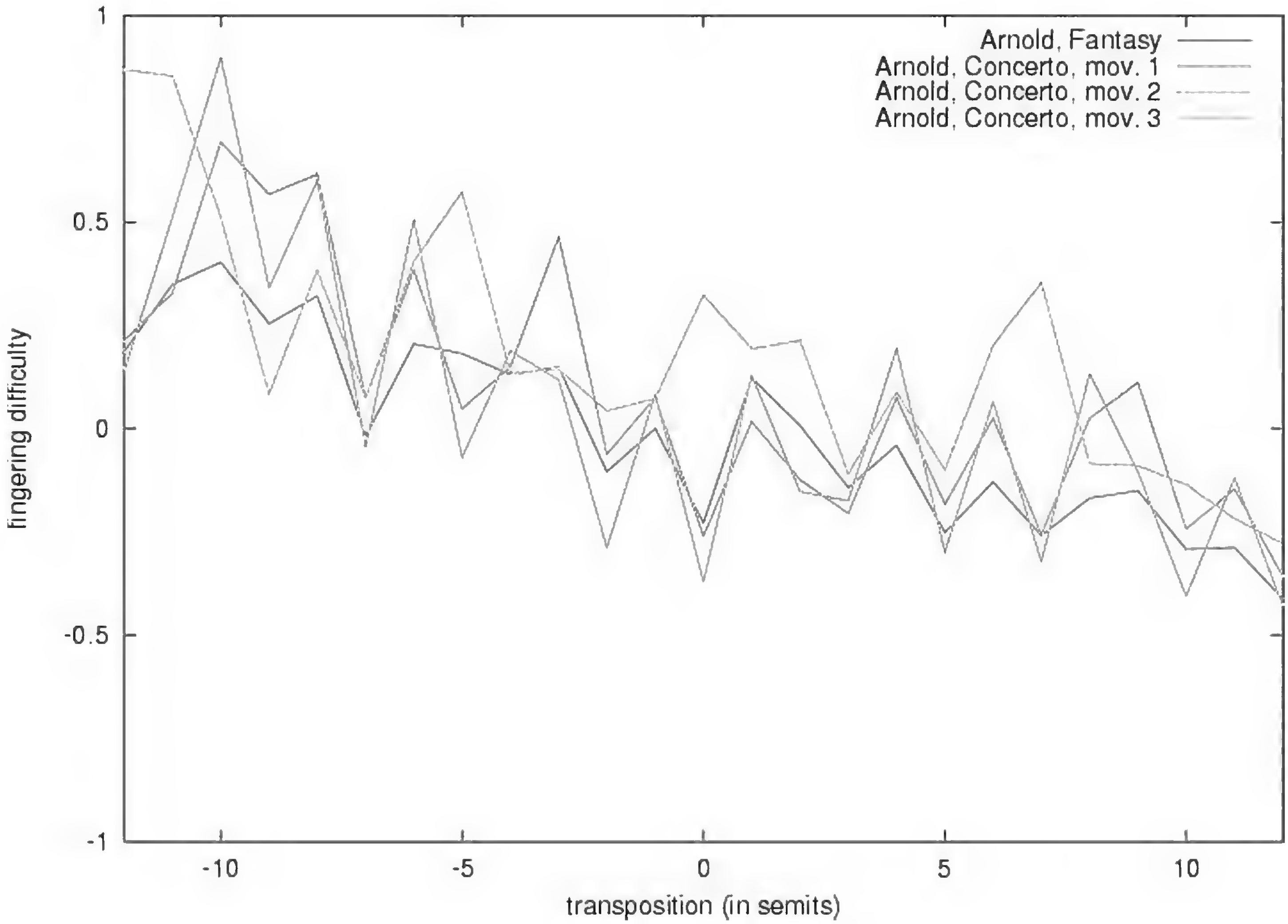


Figure 9: Effect of transposition on fingering difficulty in Malcolm Arnold's *Fantasy* and *Concerto* for trumpet.

Figure 10.

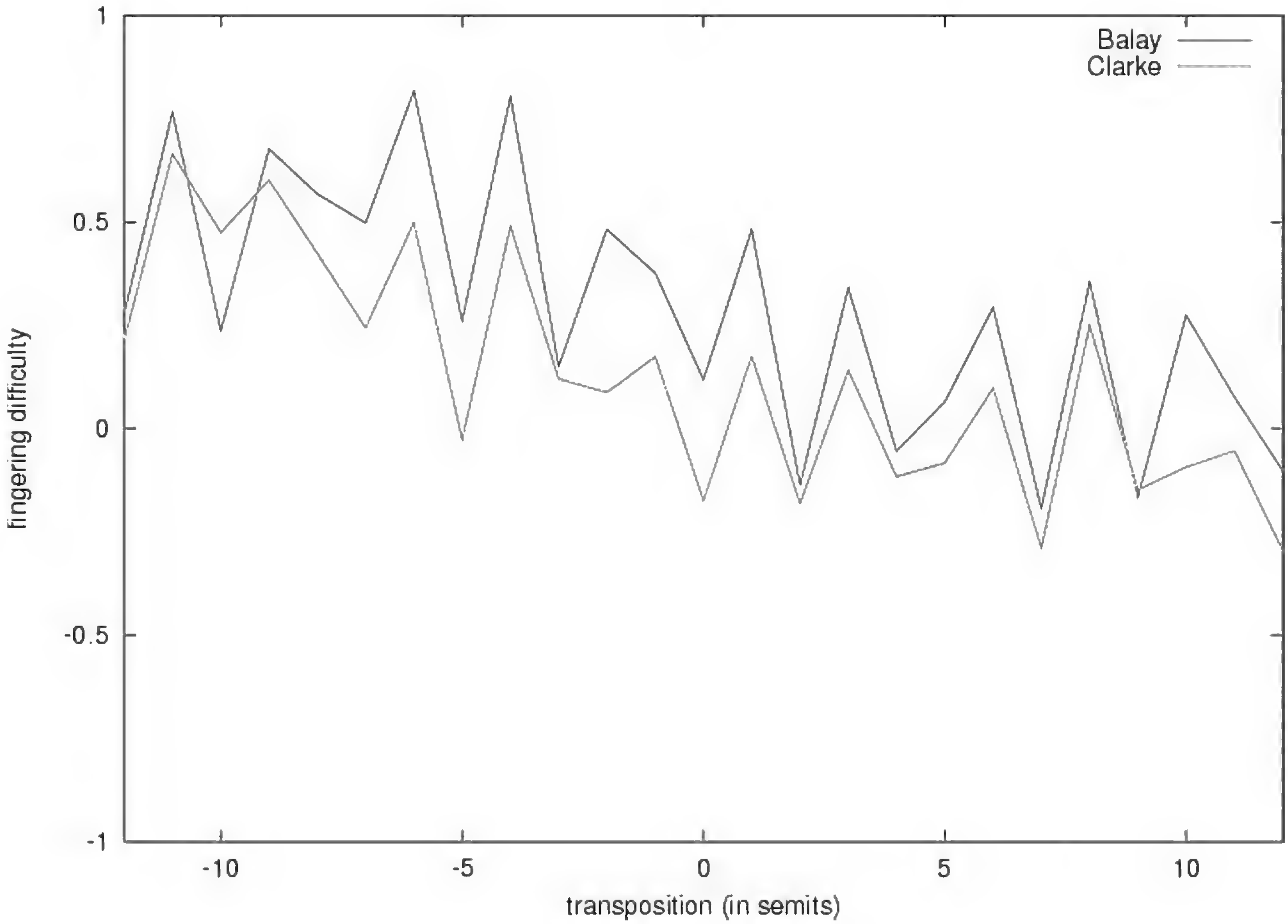


Figure 10: Effect of transposition on fingering difficulty in Guillaume Balay's *Prélude et ballade* and Herbert Clarke's *Stars in a Velvety Sky*.

Now compare these results with those for Paul Hindemith's *Trumpet Sonata* shown below. Here there is no clear effect of key, nor is there any notable dip coinciding with the key chosen by Hindemith.

Figure 11.

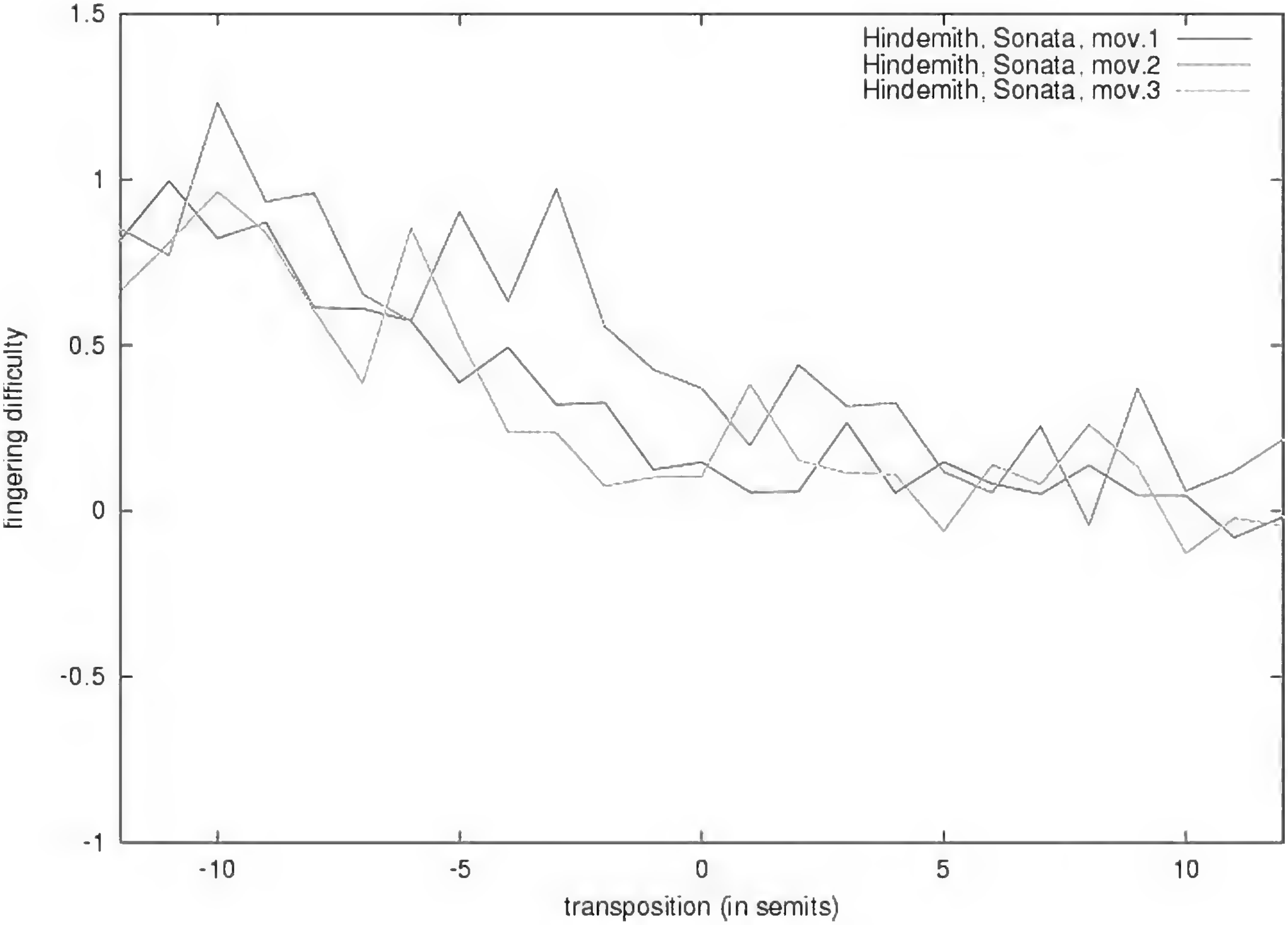


Figure 11: Effect of transposition on fingering difficulty in Paul Hindemith's *Sonate* for trumpet.

Another way to examine possible idiomatic design in these works is to observe the effect of changing the tempo. In general, as the tempo is increased, tonguing becomes more difficult while breathing becomes easier. The following graphs show the effect of tempo on overall difficulty for the works written by trumpet virtuosi. In the case of Malcolm Arnold's works, tempo seems to have little effect, except for the lively first movement of his trumpet concerto, which shows a notable increase in difficulty when the tempo is increased by roughly 25 percent.

Figure 12.

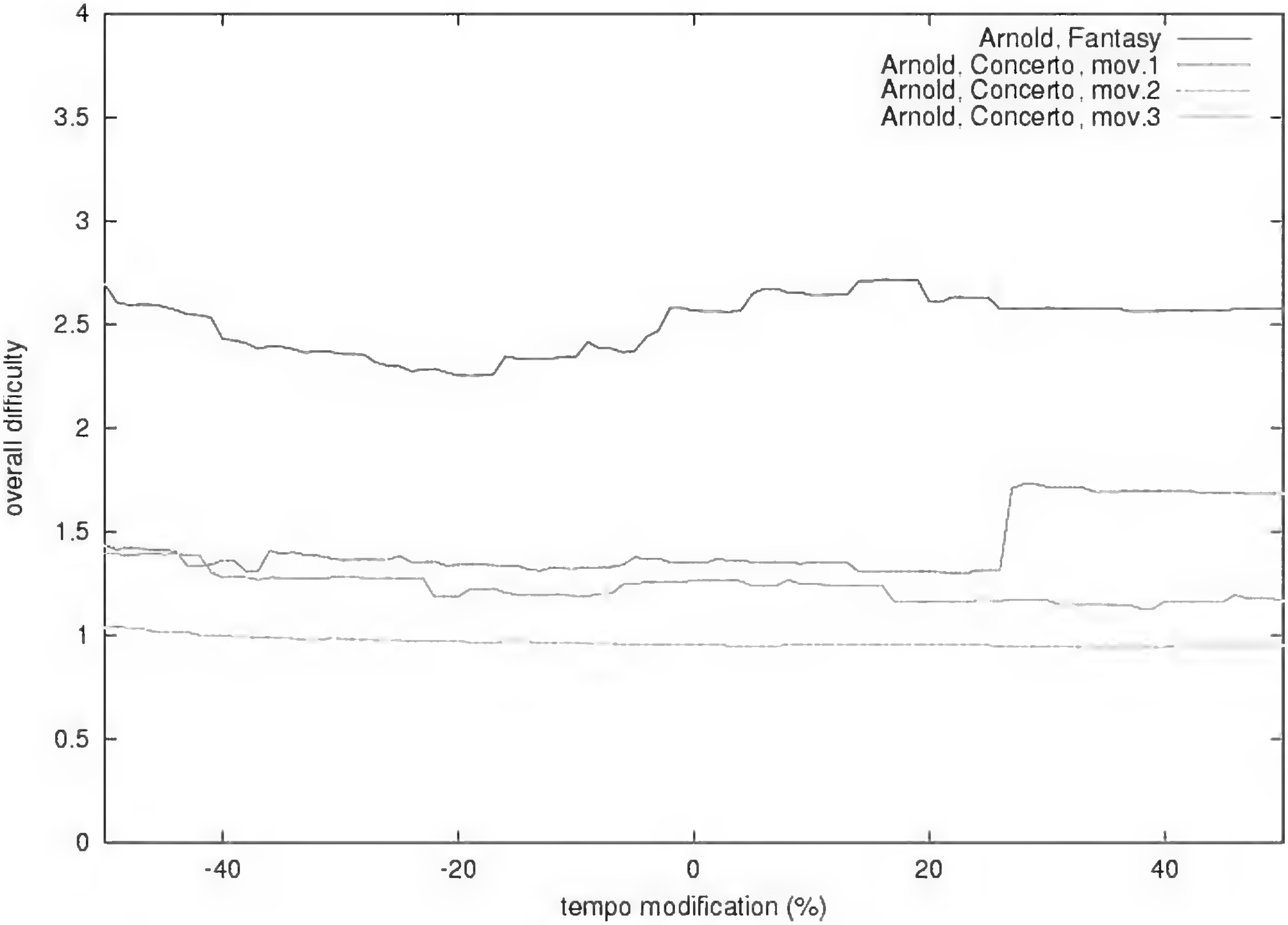


Figure 12: Effect of tempo on difficulty in Malcolm Arnold's *Fantasy* and *Concerto* for trumpet.

More dramatic changes are evident in the Balay and Clarke works, where there is a marked increase in difficulty that occurs -- a sort of "brick wall" -- where a slight increase in tempo causes a large increase in difficulty. Once again, the zero value along the X-axis corresponds to the original tempo specified by the composer in the score. Notice that for the Balay and Clarke works, the recommended tempo occurs just prior to the brick wall of increased difficulty.

Figure 13.

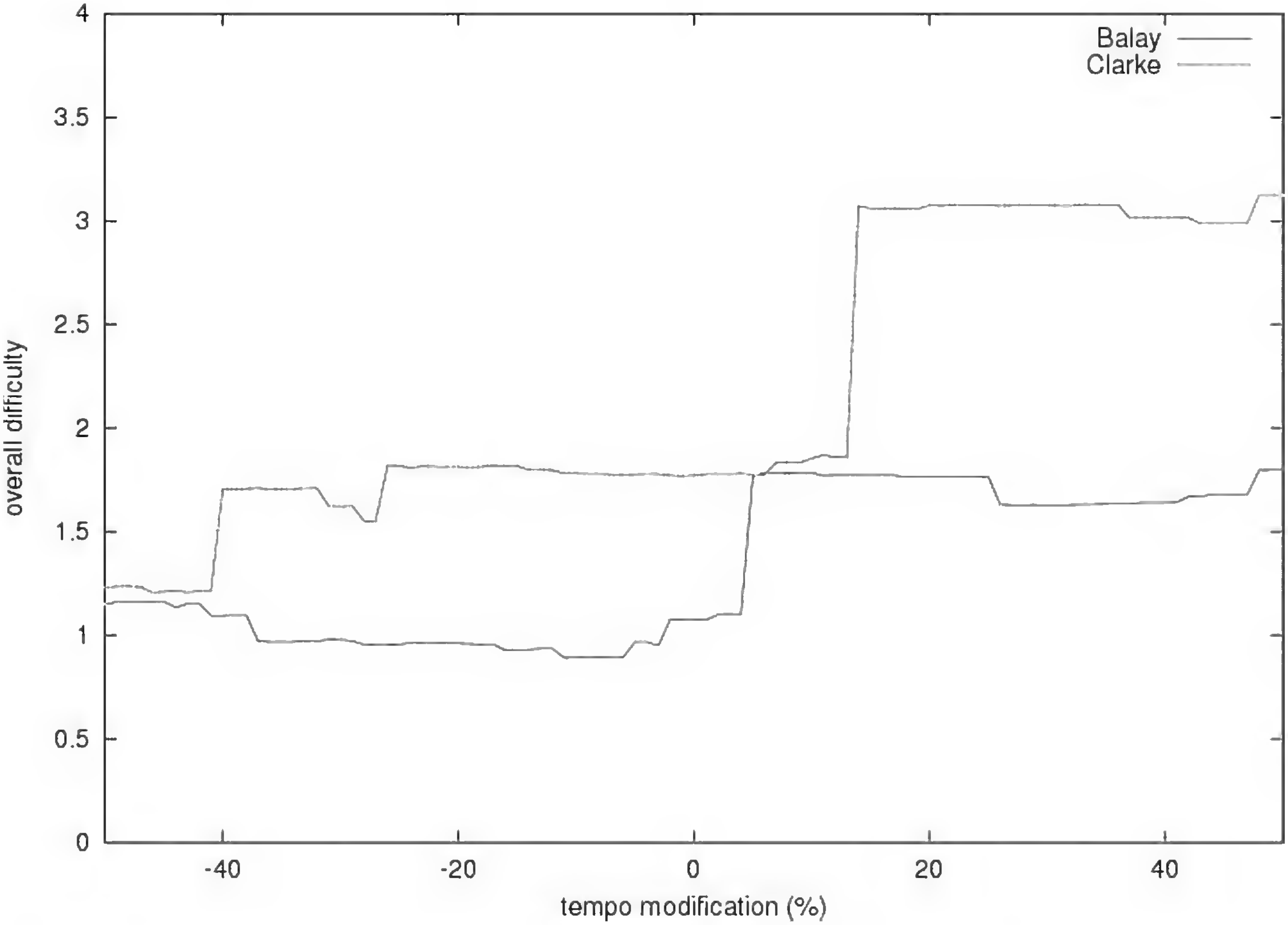


Figure 13: Effect of tempo on difficulty in Guillaume Balay's *Prélude et ballade* and Herbert Clarke's *Stars in a Velvety Sky*.

The equivalent graph for the three movements of Hindemith's *Trumpet Sonate* is shown below. By comparison with the works by trumpet virtuosi, the effect of tempo is rather featureless. In the first and third movements, the difficulty declines slightly as the tempo is increased, suggesting that the principal difficulty in these movements is linked to breathing rather than articulation.

Figure 14.

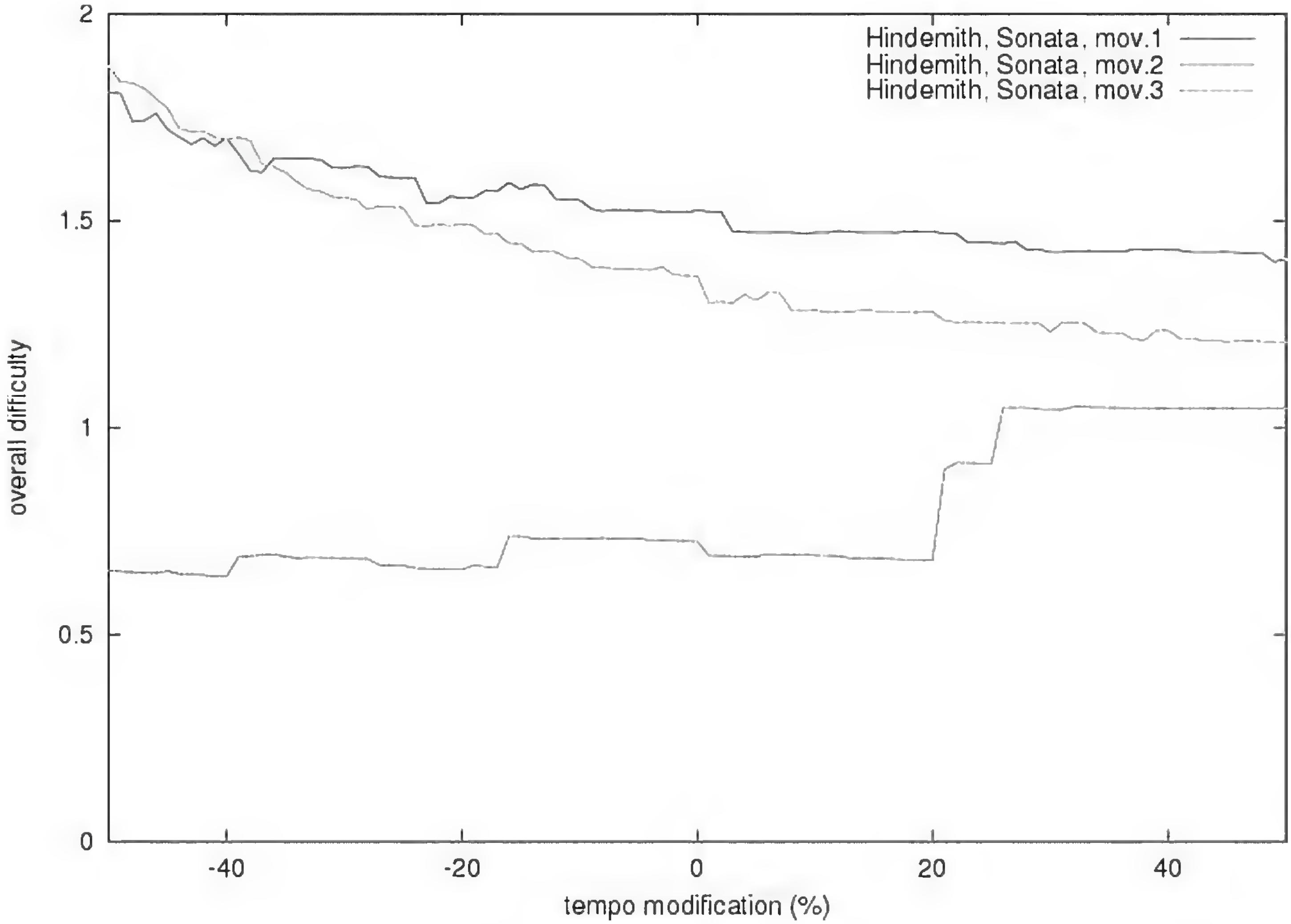


Figure 14: Effect of tempo on difficulty in Paul Hindemith's *Sonate* for trumpet.

To summarize, we've seen that the choice of key and the choice of tempo can have a considerable impact on the overall performance difficulty for a work. In the case of our sample of works by virtuoso performer/composers, we can see that the choice of keys and tempi often approach optimal values. That is, for many movements, the composer has chosen the best possible key or tempo, from the point of view of reducing the performance difficulty. In the case of a work composed by a non-trumpet player, the choice of key and tempo seems to be independent of considerations of performance difficulty.

It bears emphasizing that measures of performance ease and measures of instrumental idiomaticism cannot be regarded *prima facie* as indices of compositional merit. Difficult works are not necessarily better than easy works, and idiomatic works are not necessarily better than unidiomatic works. Only if the composer's explicit goal is to create a highly idiomatic work might such measures be construed as having a bearing on the evaluation of a composition. Moreover, there are occasionally good reasons for a composer to write explicitly difficult works. As Bernard Holland has pointed out, difficulty itself can be a handy muse.

The point of this analysis has not been to somehow denigrate Hindemith's music. Rather, my point is that musical works exhibit varying degrees of influence of the instrumental idioms. These idioms get reflected in the mental habits of performer/composers, and find their way into the very fabric of the music. That is, the performer's actions get embodied in the music. By paying close attention to the biomechanics and physiology of different performance resources, it is possible to observe idiomatic features present in the musical notation. A

virtuoso or idiomatic composer often produces works that exhibit concrete manifestations of the cognitive structures of performance.

It should be clear that we can use this approach to address analytic, historical and cognitive issues in music. For example, this approach might provide additional pertinent evidence in debates and hypotheses related to the origin of a particular work. Did composer X originally write composition Y for instrument Z, and only later arrange the work for instrument W? Finally, this approach allows us to pinpoint those aspects of musical organization that arise from the physiological, mechanical (and possibly psychological) aspects of performance.

4. Social Mediation of Taste

Idiomaticism highlights an interesting aspect of musical experience. Two instrumentalists can have very different experiences playing the same work depending on the performance situation. Yet the sonic result may be indistinguishable to the ear. For example, a difficult passage for violin might be much easier to play using a *scordatura* (re-tuning of the instrument). Of course, the same divergence of experience can also occur for listeners: two listeners hearing the same music can have dramatically different experiences. Nowhere is this phenomenon more evident than in the case of musical taste. Consider the following two examples reported by Clements:

1. A common problem for convenience stores is that they become hangouts for young teenagers. In most circumstances, the teenagers are harmless and are not breaking any law. However, store-owners regard their presence as a deterrent for other customers. It has been found that an effective strategy for minimizing loitering is to play music by the Beatles or the Beach Boys (Clements, 1993).
2. A Chicago school uses music as a punishment during after-school detentions. Detentions last 30 minutes during which the student must listen to recordings of Frank Sinatra. Students are not allowed to do homework or to talk. However, students are invited to sing along if they wish; none do (Clements, 1993). The music has made detention hall highly unpopular, and school officials are pleased by the reduced numbers of students who receive detentions.

In the first case, music has been used as a deterrent. In the second case, music is explicitly used as a punishment. What is interesting about these cases is how the popularity of the music has changed. In the 1960s, playing the music of the Beatles or the Beach Boys would probably have attracted more teenagers to loiter around the local convenience stores. Playing Frank Sinatra in the late 1950s might have made detention hall the single most popular activity at school.

What could explain the reception of music shifting from highly desirable to highly distasteful? After all, the recordings of Sinatra, the Beatles, and the Beach Boys have not changed: they are the same recordings, with the same sequences of sonic events. The music has not changed. What has changed is the people.

It is easy here to jump to conclusions about what is going on. We should acknowledge that there are several possible explanations for such dramatic changes of taste. One possibility is that modern teenagers have a different listening history. The music that has been produced since Sinatra and the Beatles has undoubtedly transformed our hearing; the music may in some sense have been superseded or have lost its power to engage or delight. This might be called the "jaded palette" hypothesis. Although a person might have loved X at one time, X is not nearly so appealing now that one can listen to Y instead.

Of course, a more popular view is to regard such changes in taste as manifestations of peer-related social interaction, especially during post-puberty years. It seems reasonable to assume that past music cannot serve to establish a distinctive peer-group identity for any new generation, since the music will continue to evoke associations with some existing age group. I will have more to say on this topic in Lecture 2 on music's origins.

At a minimum, cases such as these raise interesting questions about the representation of taste. Are musical styles and individual works mentally represented as having specific social connotations? If so, how is music represented socially?

5. Mental Representations as Brain Representations

Perhaps the ultimate representations for music are to be found in the neural codings of human brains. At the moment, we have little understanding of how the brain represents music. However, we can observe what happens when the normal representations are disrupted. Throughout history, neurologists have learned a great deal from those unfortunate individuals who have suffered physical insults to the brain.

In the area of music, Isabelle Peretz has recently written about an especially interesting case, a woman known only as "IR". IR suffered a stroke that left her with some serious musical debilitations. IR suffered no speech-related deficits, but her music listening was severely disrupted. In particular, her stroke severely damaged her musical memory. IR is not able to name well-known melodies. Moreover, she can't even identify whether a melody is familiar or unfamiliar. This is true even for very common melodies such as the national anthem. This memory deficit is evident for both long-term and short-term memory. For example, IR cannot determine whether two three-note fragments are the same or different. She can listen to an entire musical piece, and then be unable to tell whether the same piece is being played a second time.

IR can't identify violations of pitch or temporal structure, but she *can* identify violations of mode (major/minor) and tempo. She can also describe the emotional character of musical excerpts.

These deficits might not be of interest except for the following fact: IR continues to take pleasure in listening to music. Dr. Peretz gave her a cassette tape containing some music. IR enjoys playing the tape in the cassette deck in her car. She is aware that she plays the tape again and again, but each time the music is fresh and new. She enjoys the music, but cannot tell you anything about it, and cannot recognize any of the tunes from the tape when they are played.

IR raises some difficult questions for music scholars. Most theories of musical aesthetics presume that some sort of short- and medium-term memory is essential for proper musical enjoyment. But IR's listening is restricted to a paper-thin musical present in which past musical events are immediately forgotten, and future musical events remain untethered to what happened earlier.

Interactions Between Biology and Culture

As should now be clear, one of my principal concerns is to bridge the divide between those who regard music as almost exclusively cultural (with little or no influence from biology), and those who regard music as principally a sensory/perceptual phenomenon (with only a minor role for culture). It is, I believe, essential to study music from both perspectives simultaneously.

Musical phenomena are not either/or when it comes to biology and culture. Depending on the phenomenon, biology or culture may have the upper hand. In many cases, there are fascinating interactions between the two.

Let me make this claim concrete by offering an example. I'll begin by talking about an issue from a biological perspective, and then I'll look at the same issue from a cultural perspective.

In most of the world's cultures, there is a notable tendency to place the principal musical line or melody in the uppermost voice or part. This tendency is not universal; in Western music, counter examples include *faux bourdon,* barbershop quartets, and *descant* singing. Nevertheless, in general, melodies tend to be placed in the highest part.

A plausible explanation for this practice comes from what hearing scientists have discovered about auditory masking. Masking is the tendency for one sound to obscure or render inaudible another sound. Auditory masking is known to arise due to the mechanics of the basilar membrane in the cochlea, and arises when sounds are close in frequency. Two neighboring frequencies will tend to obscure each other, but the tone with the lower amplitude is prone to being completely masked.

Consider the following illustration. Suppose that two musical parts have equal amplitudes and that they both use complex tones having identical spectral content. In general, complex tones have progressively less energy in the upper partials. Figure 13 shows declining amplitudes for the first seven harmonics of a complex tone whose fundamental is 230 Hz. The X-axis has been scaled according to the position of maximum excitation along the basilar membrane; consequently, equal horizontal distances represent equal regions of potential masking. Masking will occur only between partials that are within a millimeter of each other.

Figure 13

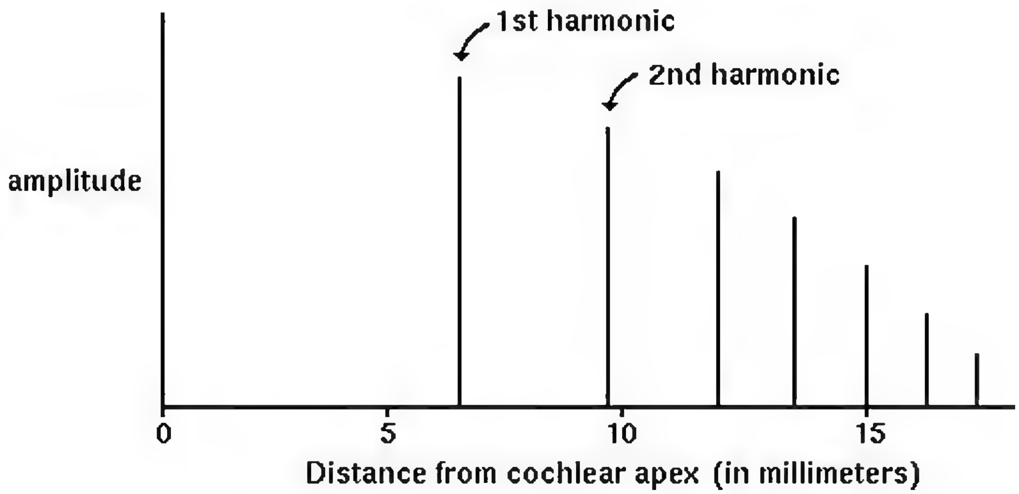


Fig. 13: Spectral content of 230 Hz complex tone.

Now consider the interaction of this tone with a 100 Hz tone having an identical spectral recipe. Partials from both tones will tend to overlap. In Figure 14, the partials of the lower tone are shown as dotted lines:

Figure 14

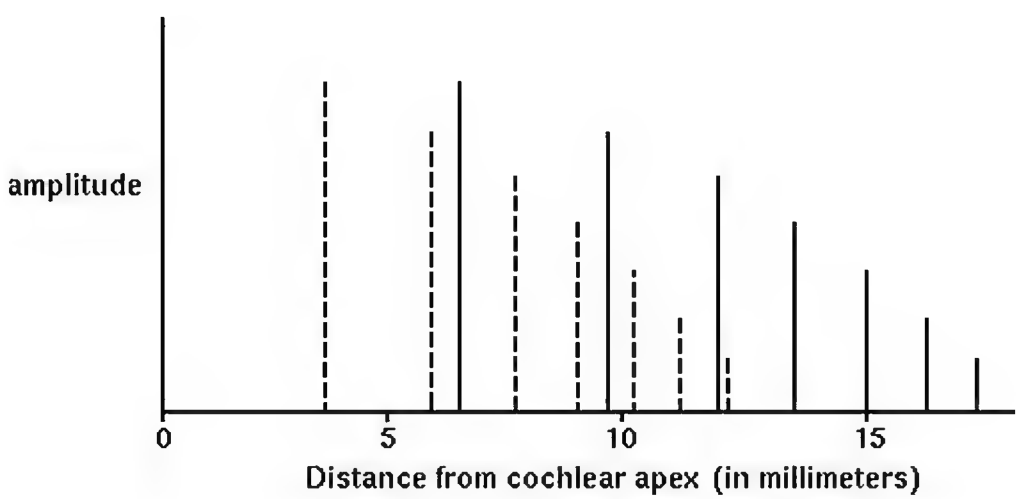


Fig. 14: Spectral interaction for two complex tones.

Notice that the upper partials of the lower-pitched tone are significantly lower in amplitude than the neighboring partials of the higher tone. Since spectral energy tends to decrease with successive partials, higher-pitched tones will tend to mask the partials of lower-pitched tones more than the reverse.

For those who understand auditory physiology, this account gives a rather satisfying explanation for why musicians might want to place the most important melodic part in the highest voice in a texture.

Let me switch gears now, and talk about one of the most robust and pervasive social phenomena attending music: namely, the long-standing and systematic discrimination against female musicians.

There is, for example, no compelling evidence (or even suggestive evidence) that women as a group are somehow inferior to men in musicianship or musical connoisseurship. Wherever women have been given an equal opportunity to pursue their musical goals, they have shown no less ability than men. However, all of the historical evidence suggests that women have been systematically sidelined when it comes to music.

It is against a background of sustained and widespread prejudice against women that the importance of auditory masking is put in perspective. In light of this prejudice, it is remarkable that so much of the music of the past would be organized to permit women to sing the foremost vocal part. Even when women were entirely excluded from music-making, it is striking that young boys (also of comparatively low social status) still managed to command the principal melodic part.[5]

We see here a complex musical phenomenon that has both biological and socio-cultural origins. In this particular case, we see a phenomenon where physiological factors mitigated an otherwise powerful social practice. The mechanics of the basilar membrane facilitated the participation of women and children in music-making. Were it not for this physiological phenomenon, one can scarcely imagine how much more profoundly women would have been excluded from the production of music.

The take-home message is not that biological factors are more important than social and cultural factors when it comes to music. (One can easily identify musical phenomena where socio-cultural factors are preeminent.) Rather, the lesson is that biological issues broadly intersect with cultural issues in intricate and interesting ways, and that a fuller understanding of music will require attention to both realms.

This lesson has been difficult to learn, not least among cognitive musicologists themselves. In her otherwise excellent book *Music As Cognition,* Mary Louise Serafine clearly expressed the formerly common view that, when it comes to music, biology is not important.

> "it is clear that the basilar membrane (or whatever structure) has exerted no appreciable influence on the way the world's music actually turned out." [p.59]

As we've seen, this isn't entirely accurate. In fact, one might be justified in claiming that, in the darkest periods of gender prejudice, it was the idiosyncracies of the basilar membrane that assured a place for women and children in music-making. Serafine's statement echoes the early attitudes in cognitive psychology when physiology and psychobiology were denigrated, primarily because of their continued association with behaviorism. Most cognitive musicologists are no longer so sanguine, and like cognitive psychologists generally, pay closer attention to the developments in cognitive neuroscience, and seek to better understand some of the biological foundations for mental activity.

Conclusion

This brings us to the conclusion of the first lecture. In this lecture I have placed cognitive musicology within the general history of the cognitive revolution. This revolution, as you will recall, arose in response to the limitations of behaviorism. The cognitive approach eschewed the positivist fallacy of interpreting absence of evidence as evidence of absence. This approach provided greater intellectual space for entertaining theories of plausible invisible mental functions. Cognitivists paid special attention to mental representations.

As we have seen, there is excellent evidence that musically pertinent mental representations exist. Ordinary listeners have access to mental representations for music, and can introspect musically. Some representations can be accessed in the total absence of sound. We can manipulate these mental representations in a variety of ways, but we cannot manipulate them in any way we wish. We've learned there is a difference between *hearing* and *hearing as,* and that scale function is a good example of the latter phenomenon. We learned that these ways of hearing are typically automatic and unconscious, and that some ways of *hearing as* are considerably easier than

others. We also saw that *hearing as* is related to culture and that the functional vocabularies are learned passively from the cultural milieu of the listener.

We've seen that listeners, even non-musician listeners, can experience passages according to rhetorical categories or types. We've noted that there exist mental habits embodied in listening styles, and that most listeners have more than one listening approach which they can apply depending on the circumstance. We've also seen evidence suggesting that the most common conscious mental activity while listening to music is daydreaming. Most of the essential aspects of music listening occur as unconscious mental processes.

We've seen that musical notations can provide useful windows to musical thought, and that modern and ancient notations can be analyzed to reveal patterns of behavior that might otherwise go unnoticed. For example, with appropriate modeling, we can see the effect of instrumental or vocal idioms on musical organization.

We've seen evidence, in the case of melodic accent, that suggests that what modern listeners hear as accented is the same as what ancient listeners heard as accented. We've seen how analyses of sound recordings point to possible social factors involved in performance practice.

We've also seen how brain injuries can sometimes give us useful clues about how mental representations are concretely coded, and how the ensuing musical changes can tell us something about the elements of musical experience. And finally, I've shown how biology and culture can interact in subtle and unexpected ways -- as when the structure of the human hearing organ tended to mitigate against a pervasive sexism.

The Promise of Cognitive Musicology

What is cognitive musicology? Cognitive musicology is the study of *habits of mind* as they relate to music. Since minds are the products of both biology and culture, cognitive musicology is an approach to the study of music that takes both biology and culture seriously. A common ground for both biological and cultural study is found in the domain of mental representations. Consequently, much of the day-to-day research of cognitive musicologists centers on discovering and deciphering various music-related mental representations.

As you might expect, I believe that cognitive musicology has much to offer music scholarship in general.

For the historian, cognitive musicology offers (with some limitations) the possibility of reconstructing aspects of seemingly lost practices. It also offers ways to approach how musical works and practices may have held meanings for listeners and musicians of past historical periods and places.

For the ethnomusicologist, cognitive musicology offers relatively effective techniques for gaining access to the minds of others, and useful ways of pinpointing how culturally sophisticated experiences differ from culturally naive experiences. Cognitive musicology also offers the ethnomusicologist better ways for investigating how material and cultural conditions get reflected and expressed in a music.

For the performer, cognitive musicology offers ways for investigating what distinguishes inexpressive and pedestrian performances from inspired and compelling ones.

For the composer, cognitive musicology offers pointers to cognitively and perceptually rich regions of unexplored musical materials. In describing musical "habits of mind," cognitive musicology can help composers in their quests to establish *new* habits for the musical mind.

For the music theorist, cognitive musicology promises to address basic questions of musical organization from a more rigorous and less speculative approach.

There has been a growing interest in music cognition in recent years. I think this growth originates, at least in part, because cognitive musicology can appeal to scholars inspired both by continental and by Anglo-American philosophical traditions. For the continentally-inspired scholar, music cognition offers the opportunity to treat

subjectivity as real without reifying it. Music cognition provides ways of considering the subjective without making it mystical or juxtaposing it irredeemably against the objective. Nor does it merely objectify the subjective.

For the empiricist-inspired scholar, cognitive musicology offers the opportunity to transform intuition and speculation into conjecture and hypothesis, and thereby provides a means for testing musical ideas and theories.

In the ensuing lectures, I hope to illustrate in greater detail some of the accomplishments and opportunities that cognitive musicology holds.

Thank you.

Footnotes

[1] At the same time, those music scholars who pursued socially-oriented studies in music (such as the Anglo-Marxist popular-music scholars) failed to pay much heed to the extant psychological research. As the anthropologist Roy D'Andrade has pointed out (regarding sociology generally), sociologists have showed an extraordinary ignorance of the extant psychological research, and have tended to devise their own psychological theories with little reference to the existing research.

[2] In an unguarded moment, Ulric Neisser unhelpfully wrote that "every psychological phenomena is a cognitive phenomena." This casts a very wide net. As we will see, there are a number of themes that characterize and give some focus to cognitive psychology and cognitive science.

[3] This enthusiasm was concretely evident in research on information processing, where mental phenomena were analyzed as successively ordered stages of processing.* [See, e.g., R. Lachman, J. Lachman & E.C. Butterfield, *Cognitive Psychology and Information Processing.*] As Ulric Neisser defined it,

> "Cognitive psychology refers to all processes by which the sensory input is transformed, reduced, elaborated, stored, recovered, and used."

[4] Examples of other principles might include (1) phrase-final fall (where pitches at the ends of phrases *tend* to exhibit a downward contour), (2) a preponderance of small intervals; in particular, repeated pitches are common (except when it is impossible to re-articulate notes -- such as with the bagpipe) (3) repeated text is often associated with repeated melodic passages (facilitates memory)

[5] There may be other factors that also favor placing the melody in the upper-most voice. However, auditory masking appears to play the most significant role.

References

Arnold, M. (1969).
Fantasy for B flat Trumpet. (Opus 100), London: Faber Music Ltd.

Baily, J. (1985).
Musical structure and human movement. In: I. Cross & R. West (Eds.), *Musical Structure and Cognition.* London: Academic Press, pp. 237-258.

Balay, G. (n.d.)
Prélude et ballade. Cornet Solo with Piano Accompaniment. New York: Belwin Inc.

The Benedictines of Solesmes (Eds.). (1963).
The Liber Usualis. Tournai, Belgium: Descleee Company.

Bowen, J.A. (1993-4)
A computer aided study of conducting. *Computing in Musicology,* 9: 93-103.

Bruner, C.L. (1984).
The perception of contemporary pitch structures. *Music Perception,* 2(1), 25-39.

Caplin, W. (1978).
Der Akzent des Anfangs: Zur Theorie des musikalischen Taktes. *Zeitschrift für Musiktheorie,* 8, 17-28.

Clarke, E.F. (1989).
Mind the Gap: Formal structures and psychological processes in music. *Contemporary Music Review,* 3, 1-13.

Clarke, H.L. (n.d.)
Stars in a Velvety Sky. Solo B-flat cornet. New York: Carl Fischer.

Collins, D. & Huron, D. (in press).
Voice-leading in Cantus Firmus-based canonic composition: A comparison between theory and practice in renaissance and baroque music. *Computers in Music Research.*

Densmore, F. (1929).
Pawnee Music. Washington, DC: Smithsonian Institution, Bureau of American Ethnology.

Drake, C.W., Dowling, W.J. & Palmer, C. (1991).
Accent structures in the reproduction of simple tunes by children and adult pianists. *Music Perception,* 8, 315-334.

Farnsworth, P.R. (1948).
Sacred cows in the psychology of music, *Journal of Aesthetics and Art Criticism,* 7(1), 48-51.

Francès, R. (1958/1988).
La Perception de la Musique. Translated by W.J. Dowling as *The Perception of Music.* Hillsdale, N.J.: Lawrence Erlbaum Associates.

Gibson, D. (1986).
The aural perception of nontraditional chords in selected theoretical relationships: A computer-generated experiment. *Journal of Research in Music Education,* 34(1), 5-23.

Gibson, D. (1988).
The aural perception of similarity in nontraditional chords related by octave equivalence. *Journal of Research in Music Education,* 36(1), 5-17.

Gibson, D. (1993).
The effects of pitch and pitch-class content on the aural perception of dissimilarity in complementary hexachords. *Psychomusicology,* 12(1), 58-72.

Harwood, D.L. (1976).
Universals in music: a perspective from cognitive psychology. *Ethnomusicology,* 20, 521-533.

Hindemith, P. (1940).
Sonate für Trompet in B und Klavier. Mainz: B. Schott's Sohne (ED 3643).

Holland, B. (1999).
When composers make it hard, fright and strain become muses. *New York Times,* June 1, 1999; p.B1f.

Huron, D. (1990).
Mary Louise Serafine: Music As Cognition: The Development of Thought in Sound [review of]. *Psychology of Music,* Vol. 18, No. 1, pp. 99-103.

Huron, D. & Berec, J. (1993).
The influence of performance physiology on musical organization: A case study of idiomaticism and the B-flat valve trumpet. Unpublished manuscript.

Huron, D. & Royal, M. (1996).
What is melodic accent? Converging evidence from musical practice. *Music Perception,* 13(4), 489-516.

Kippen, J. & Bell, B. (1989).
The identification and modelling of a percussion "language", and the emergence of musical concepts in a machine-learning experimental set-up. *Computers & Humanities,* 23(3), 199-214.

Krumhansl, C.L. (1990).
Cognitive Foundations of Musical Pitch. Oxford: Oxford University Press.

Krumhansl, C.L. (1995).
Music psychology and music theory: Problems and prospects. *Music Theory Spectrum* 17(1), 53-80.

Lannoy, C. (1972).
Detection and discrimination of dodecaphonic series. *Interface,* 1, 13-27.

Largent, E.J. (1972).
An investigation into the perceptibility of twelve-tone rows. Ohio State University, unpublished PhD Dissertation.

Lewin, D. (1986).
Music theory, phenomenology, and modes of perception. *Music Perception,* 3(4), 327-392.

Lomax, A. (1962)
Song structure and social structure. *Ethnology,* 1, 425-451.

MacKenzie, C.L. & Iberall, T. (1994).
The Grasping Hand. Amsterdam: North-Holland (Elsevier Science).

Millar, J.K. (1984).
The aural perception of pitch-class set relations: A computer-assisted investigation. North Texas State University, unpublished PhD dissertation.

Nam, U. (1998).
Pitch distributions in Korean Court music: Evidence consistent with tonal hierarchies. *Music Perception,* 16(2), 243-248.

Neisser, U. (1967).
Cognitive Psychology. New York: Appleton-Century-Crofts.

Pedersen, P.R. (1970).
The perception of musical pitch structure. University of Toronto, unpublished PhD Dissertation.

Roederer, J.G. (1987).
Why do we love music? A search for the survival value of music. In: *Music in Medicine,* R. Spintge & R. Droh (Eds.), Heidelberg: Springer Verlag.

Serafine, M.L. (1988).
Music As Cognition: The Development of Thought in Sound. New York: Columbia University Press.

Squire, C.R. (1901).
Genetic study of rhythm. *American Journal of Psychology,* 12, 546-560.

Sudnow, D. (1978).
The Ways of the Hand; The Organization of Improvised Conduct. Cambridge, Massachusetts: Harvard University Press.

Sudnow, D. (1979).
Talk's Body. New York: Alfred A. Knopf, Inc.

Temperley, D. (2000).
The question of purpose in music theory: Description, suggestion, and explanation. *Current Musicology,* in press.

Thomassen, J. (1982).
Melodic accent: Experiments and a tentative model. *Journal of the Acoustical Society of America,* 71, 1596-1605.

Thomassen, J. (1983).
Erratum. *Journal of the Acoustical Society of America,* 73, 373.

Thrall, B. (1962).
The audibility of twelve-tone serial structure. Ohio State University, unpublished PhD dissertation.

Tuzin, D. (1984).
Miraculous voices: the auditory experience of numinous objects. *Current Anthropology,* 25, 579-596.

Vos, P.G. & Troost, J.M. (1989).
Ascending and descending melodic intervals: statistical findings and their perceptual relevance. *Music Perception,* 6(4), 383-396.

Watt, H.J. (1924).
Functions of the size of interval in the songs of Schubert and of the Chippewa and Sioux Indians. *British Journal of Psychology,* 14, 370-386.

Weber, M. (1958).
The Rational and Social Foundations of Music. Carbondale: Southern Illinois University Press.

Woodrow, H. (1911).
The role of pitch in rhythm. *Psychological Review,* 18, 54-77.

Yung, Bell (1984).
Choreographic and kinesthetic elements in performance on the Chinese seven-string zither. *Ethnomusicology,* Vol. 28, pp. 505-517.

University of California, Berkeley
Department of Music

A shorter version of this article can be found in: D. Huron (2001), "Is music an evolutionary adaptation?" Annals of the New York Academy of Sciences, *Vol. 930, pp. 43-61.*

The 1999 Ernest Bloch Lectures

Lecture 2. An Instinct for Music:

Is Music an Evolutionary Adaptation?

David Huron

Addressing the question of music's origins has a long -- though not particularly distinguished -- history. Many cultures have provided colorful stories describing how humans acquired the capacity for music. A few bold or wreckless scholars have ventured to offer biological, psychological, social, cultural or religious accounts of possible origins. Most scholars have wisely steered clear of the issue of music's origins, since clearly the enterprise is patently speculative. At its worst, proposals concerning musical origins are fiction masquerading as scholarship. However, I think there remains some merit in contemplating the question of how music-making got started. Reflecting on such questions can be a potentially informative and perhaps even illuminating exercise. In this lecture, I propose to offer a social account of music's origins that is explicitly linked to one of the most successful theories yet devised: the theory of evolution by natural selection.

Richard Dawkins (1995) reminds us of the importance of natural selection in the following passage:

> "All organisms that have ever lived -- every animal and plant, all bacteria and all fungi, every creeping thing ... can look back at their ancestors and make the following proud claim: Not a single one of our ancestors died in infancy. They all reached adulthood, and every single one [successfully reproduced.] Not a single one of our ancestors was felled by an enemy, or by a virus, or by a misjudged footstep on a cliff edge, before bringing at least one child into the world. Thousands of our ancestors' contemporaries failed in all these respects, but not a single solitary one of our ancestors failed in any of them. These statements are blindingly obvious, yet from them much follows: much that is curious and unexpected, much that explains and much that astonishes." (1995; p.2)

The theory of evolution is possibly the most powerful theory yet conceived. It has survived the strongest of challenges, such as the challenge of accounting for altruistic behavior. With the work of Hamilton, Trivers, Wilson, and others, and the extraordinary accomplishments of molecular geneticists, the theory of evolution has gone from strength to strength. If Darwin were alive today, he would be very impressed by just how many phenomena appear to be accounted for by his theoretical legacy.

Evolution is often thought of in purely physiological rather than psychological terms. It is not simply that evolution has shaped immune systems, digestive tracts, and knee caps. Evolution has also shaped our attitudes, dispositions, emotions, perceptions, and cognitive functions. Some of our deepest convictions can be traced to

plausible evolutionary origins: we love life, we fear death, and we nurture our children because any group that did not have these dispositions would be at a competitive disadvantage.

In addition to psychological dispositions and attitudes, our cognitive and perceptual capacities are also the products of evolution. Cognitive and perceptual capacities are shaped by and adapted to the world. Clearly, our perceptions of the world are not precise characterizations of how the world really is. But neither are our perceptions arbitrary constructions, for if so, we would soon be dead. In the case of sound, our ways of perceiving and apprehending are first of all conditioned by how sounds behave in the physical world, and by what information sounds encode that might be of value to human survival and procreation.

Similarly, our emotional lives are shaped by evolution. Research by Randolf Nesse has shown that even sadness can serve an essential evolutionary purpose; feeling bad may not be so bad after all. Like pain, feeling bad may be unpleasant, but it may also be biologically useful.

The theory of evolution by natural selection is a *distal* theory rather than a *medial* or *proximal* theory. It is not a theory that explains specific behaviors, such as why you chose to cook ravioli for dinner last night, why you parked in a particular parking spot this morning, or why you decided to learn to play the viola. Evolution proceeds by selecting traits that are adaptive to an organism's environment. For example, evolution did not "originate" or "create" the phenomenon of *altruism*. Instead, given a certain environment, natural selection favored individuals who exhibited altruistic traits. Evolution does not dictate our behavior: it selects which behaviors are likely to be passed on to subsequent generations -- and it selects only those behaviors that have a genetic basis.

Does Music Have Survival Value?

Because so many human behaviors are clearly linked to survival, we might entertain the question of whether musical behaviors also confer some sort of advantage that enhances human survival and procreation. This is a difficult, contentious, and unresolved question. In this lecture, I will not attempt to somehow prove that music is adaptive; rather, my goal here is to convince you that this is a worthwhile question and deserves further thought and discussion.

Many knowledgeable people have concluded that music has no survival value. Indeed, a number of esthetic philosophers have argued that an essential, defining characteristic of the arts is that they serve no practical function. Accordingly, any music that is created for biological (or economic) reasons cannot be considered art. Even among evolutionary psychologists, it has been common to suppose that music is not adaptive. In *How the Mind Works,* for example, the noted psycholinguist and evolutionary psychologist Steven Pinker has argued that music is a good example of a common human phenomenon that is likely *not* an evolutionary adaptation.

I think that the evidence one way or another is not particularly convincing. Like others, I am not at all convinced that music is an evolutionary adaptation. However, I think we should investigate matters further before we dismiss the notion out of hand.

Before addressing the question of possible evolutionary origins for music, it is useful to consider some of the dangers associated with forming an evolutionary argument. Not all of the pertinent dangers can be noted here, but let me at least identify six of the more important ones.

1. In his well-known work on epistemology, *The Logic of Scientific Discovery,* Karl Popper (1935/1959) argued that the theory of evolution by natural selection lacks a scientific status because the theory as a whole cannot be directly falsified. No scientist has formulated the theory in such a way that a set of observations could, in principal, be used to falsify it. Popper consequently referred to the theory of evolution as a *pre-scientific* theory. This "non-scientific" status did not significantly diminish the importance of the theory in Popper's eyes. Popper argued that the theory remains scientifically important because of its hypothesis-generating capacity. Individual hypotheses arising from the theory of evolution (such as the Trivers-Willard hypothesis discussed below) are themselves testable.

2. A second problem associated with evolutionary reasoning is the problem of *post hoc* reasoning. Gould and Lewontin (1979) have noted that it is relatively easy to concoct theories to explain pre-existing data. For example, since we already know that camels have humps, we can generate all sorts of plausible explanations as to their origin. Like Rudyard Kipling's *Just So* stories, there are innumerable opportunities for unfounded "story-telling" (Lewontin, 1991). Philosophers refer to after-the-fact theories as "*post hoc* theories." *Post hoc* theories are properly regarded as inferior because they use the facts *twice*: first as a basis for formulating the theory, and second as "evidence" in support of that theory. Good theories, by contrast, are *a priori*; that is, the theory suggests or predicts certain facts or phenomena before these facts are ascertained or observed.

   It should be noted, however, that *post hoc* theories can sometimes develop into *a priori* theories. The transformation of a *post hoc* theory into an *a priori* theory occurs when some unexpected prediction is seen to be a logical outcome of the theory. (In many cases, such *a priori* formulations are also, in principal, falsifiable, so these theories also become "scientific" in Popper's terminology.)

   As Tooby and Cosmides (1992) have pointed out, Lewontin and Gould's critique of evolutionary reasoning is too sweeping. Although many evolutionary accounts are clearly *post hoc*, a large number of evolutionary accounts are *a priori*. For example, evolutionary theory has led to remarkably abstruse and counterintuitive predictions such as the Trivers-Willard hypothesis (Trivers and Willard, 1973). One prediction arising from this hypothesis is that human male offspring will be breast-fed longer than female offspring by mothers from high socio-economic backgrounds, while female offspring will be breast-fed longer than male offspring by mothers from low socio-economic backgrounds. In a study of North American families, this and related predictions have been confirmed (Gaulin and Robbins, 1991). Other tests have similarly proved to be consistent with predictions from the Trivers-Willard hypothesis (see Ridley, 1994, and Wright, 1994, for reviews).
3. An important issue is how we interpret the repercussions of naturalistic accounts of phenomena. Philosophers refer to the belief that `the way things are in nature is the way they ought to be' as the *naturalist fallacy*. This view conflates what *is* with what *ought to be*. The naturalist fallacy is a sort of double-edged sword, however. Whereas we properly blame sexists for failing to recognize the is-ought distinction, we don't typically blame environmentalists (for example) for their reliance on this same mode of argument. We tend to be attracted to "natural" accounts that support our views and refute the views of others. Yet when others use "nature" to support their views, we point to the naturalist fallacy. Most of us are rampant hypocrites when it comes to the naturalist fallacy. Moreover, not all philosophers are convinced it is a fallacy.

   It is entirely legitimate to be suspicious of anyone purporting to investigate possible evolutionary origins for music. Our fear is that some people will be tempted to use this information as a way of buttressing arguments concerning musical taste: music X is more natural, and therefore superior to music Y. However, I think these suspicions are overblown. Whatever the origins of music, the vast majority of people have long ceased to live in Paleolithic conditions. It is doubtful that reconstructing the sounds of neolithic caves will be more satisfying than Beethoven or Buddy Holly.
4. Evolutionary theory has been used to defend all sorts of nefarious ideologies, from racism to sexism. There is a voluminous and distinguished literature on genetic diversity and human equality, which we won't review for reasons of space. However, this literature provides important guidelines for interpreting how evolutionary arguments carry over to moral and esthetic discourse (see for example Dobzhansky, 1973). The fact that a theory may be used to support nefarious moral ideologies does not make the theory false; rather it establishes that we need to be vigilant about how theories are interpreted.
5. By discussing biological issues, an author runs the risk of being misconstrued as believing that cultural factors are unimportant. As I emphasized in my first lecture, minds are the product of both biology and culture. Like most other music scholars, I believe that culture is the more important factor. However, our belief in the preeminence of culture does not give us license to dismiss possible biological foundations.
6. If music is an evolutionary adaptation, then it is likely to have a complex genesis. Any musical adaptation is likely to be built on several other adaptations that might be described as pre-musical or proto-musical. Moreover, the nebulous rubric "music" may represent several adaptations, and these adaptations may

involve complex co-evolutionary patterns with culture (see Durham, 1991). In biological matters, things are rarely straightforward.

Given these possible dangers, why bother attempting to formulate an evolutionary theory of music? Isn't it premature? First, as noted above, my goal here is not to convince you that music is adaptive; my goal is only to convince you that this is a worthwhile question. Understanding the possible origins of music might help inform us about some of the reasons we tend to respond in certain ways. Second, in the spirit of Popper, I will aim to tell an evolutionary story that is able to generate testable hypotheses. Like other evolutionary accounts, my theory will draw on existing knowledge, and so be *post hoc* in character. As long as this account remains *post hoc*, Gould and Lewontin's criticisms raise justified and paramount difficulties. However, it is my hope that the theory can be developed to the point where testable hypotheses might be derived.

Before entertaining some possible evolutionary views of music's origins, let us first consider two pertinent complicating points of views. One view is that music is a form of non-adaptive pleasure seeking. A second view is that music is an evolutionary vestige.

NAPS Theory of Music

Most pleasurable activities, such as eating and sex, have clear links to survival. Such activities ultimately stimulate brain mechanisms that are specifically evolved to reward and encourage adaptive behaviors. Note that once brain mechanisms are in place that permit the experience of pleasure, it may be possible to stimulate those mechanisms in ways that don't confer a survival advantage. We can call these behaviors *non-adaptive pleasure-seeking* (NAPS). An example of NAPS behavior is found in the human taste for sugars and fats. In pre-modern times, sugars and fats were rare in human diets, but highly nutritious in the amounts available. There are good reasons why human tastes would evolve to reward the ingestion of foods with high fat or sugar content. However, centuries of human ingenuity have succeeded in generating a modern diet that contains unnaturally high levels of fats and sugars -- levels so high as to cause health problems such as diabetes and heart disease. Although such tastes originally conferred an increased chance of survival, in the modern environment, these behaviors have become less adaptive.

Another example of NAPS behavior is found in drug use such as heroin or cocaine. These drugs can directly activate the brain's pleasure centers, simply by injecting or imbibing a substance. Although the channel for pleasure exists for good evolutionary reasons, it may be possible to exploit the channel without any concomitant survival-enhancing result.

As in the case of drugs, it is possible that musical behaviors are forms of non-adaptive pleasure-seeking. That is, music itself may not enhance human survival; music may merely exploit one or more existing pleasure channels that evolved to reinforce some other adaptive behavior(s). We might call this view the "NAPS Theory of Music."

One way to determine whether some pleasure-seeking behavior is adaptive or non-adaptive is to consider how long the behavior has been around. In the long span of evolutionary history, non-adaptive pleasure-seeking behaviors tend to be short-lived. For example, heroin users tend to neglect their health and are known to have high mortality rates. Furthermore, heroin users tend to neglect their offspring -- they make poor parents. Poor health and neglect of offspring are infallible ways of reducing the probability that one's genes will be present in a future gene pool. After many generations, natural selection will tend to mitigate against heroin use. Those individuals who are not disposed (for whatever reason) to use heroin, are much more likely to procreate and so pass along their aversion to the use of such drugs, provided that the aversive behavior is somehow linked to a gene or genes.

The use of alcohol already suggests how NAPS behaviors can transform a gene pool. Although no gene has yet been identified, either for alcohol susceptibility or for alcohol tolerance, the responses of different human populations to alcohol show a suggestive pattern. Large quantities of alcohol became possible only with the advent of agriculture. European and Asian descendents of early agrarian cultures (such as originated in Mesopotamia) manage to deal with alcohol better than descendents of traditional hunter-gatherer societies, such

as indigenous peoples in the Americas and in the arctic regions of Europe. Of course there are certain to be non-genetic factors influencing alcohol tolerance and abuse. However, alcohol researchers suspect that genetic factors are at work. Those people who have descended from traditional agricultural societies have a clear statistical advantage in dealing with the non-adaptive consequences of alcohol, and this would be expected if alcohol had been prevalent in these societies for thousands of years.

If music itself has no survival value (and merely exploits an existing pleasure channel) then any disposition towards musical behaviors would tend to *worsen* one's survival. Spending inordinate amounts of resources (such as time and money) on music might be expected to place music-lovers at an evolutionary disadvantage. In other words, if the NAPS Theory of Music is true, then we might predict that music appreciation would be correlated with marginal existence: as in the case of alcohol, people on "skid row" might be expected to be disproportionately music enthusiasts.

If music is non-adaptive, then the likelihood is that music is a modern invention; otherwise music-lovers would have become extinct some time ago. As we will see, the archaeological evidence indicates that music is very old -- much older than agriculture -- and this great antiquity is inconsistent with music originating as a non-adaptive pleasure-seeking behavior. In short, there is little evidence that musical behaviors have been selected against. All of this suggests that there is little support for the NAPS Theory of Music.

Music as an Evolutionary Vestige

Another view might be that, while music at one time did indeed confer some survival value, it is now merely vestigial. Like the human appendix, at one time this "organ" may have contributed directly to human survival, but now it is largely irrelevant -- an evolutionary leftover. If this view is true, then we would have to ask `What advantage *did* music once confer?' and `How have things changed so that music is no longer adaptive?'

Measuring the Adaptive Value of Music

The adaptive value of some function is often evident in the individual survival costs arising from that function. For example, the larynx of newborn infants is anatomically arranged so that breathing and swallowing can happen at the same time. When the larynx enlarges, our physiological capacity for speech is purchased at the price of the danger of choking. In fact, one measure of the evolutionary advantage of speech is the mortality rates due to choking.

Similarly, an estimate of the evolutionary advantage conferred by music is to measure the amount of time people spend in musical behaviors. In the Atlas mountains of Morocco, full-time *Jujuka* mountain musicians are supported by the local villagers. That is, there is an entire caste of people whose principal productive activity is music-making. A ready index of the importance of music in such a society may be the ratio of the number of musicians to the number of farmers and herders.

Some Evolutionary Theories of Music

Now let's consider some possible positive answers concerning the evolutionary advantage of music.

1. **Mate Selection.** In the same way that some animals find colorful or ostentatious mates attractive, music-making may have arisen as a courtship behavior. The ability to sing well might imply that the individual is in good health.
2. **Social Cohesion.** Music might create or maintain social cohesion. It may contribute to group solidarity and so increase the effectiveness of collective actions.
3. **Group Effort.** More specifically, music might contribute to the coordination of group work, such as pulling a heavy object, defending against a predator, or attacking a rival clan.

4. **Auditory Development.** Listening to music might provide a sort of "exercise" for hearing. Music might somehow teach people to be more perceptive. Similarly, music-making might provide opportunities for developing more refined motor coordination.
5. **Conflict Reduction.** In comparison with speech, music might reduce conflict. Sitting around a fire talking may well lead to arguments and possible fights. Sitting around a fire singing might provide a safer social activity.
6. **Safe Time-Passing.** Evolutionary biologists have noted that the amount of sleep an animal requires is proportional to the effectiveness of food gathering. Efficient hunters (such as lions) spend a great deal of time sleeping. Grazing animals, by contrast, sleep relatively little since they must eat for long periods each day. One argument is that sleep helps to keep an animal out of trouble. A lion is more apt to injure itself if it is engaged in unnecessary activities. A parallel argument in music might be that music provides a safe way to pass time. As early humans became more effective at gathering food, music might have arisen as a harmless pastime. (Note, for example, that humans sleep more than other primates.)
7. **Transgenerational Communication.** Given the ubiquity of folk ballads and epics, music might have originated as a mnemonic conveyance for useful information. Music might have provided a comparatively good channel of communication over long periods of time.

Sexual Selection

Before continuing, we should take a moment to discuss a common, though questionable theory of music's origins. Charles Darwin identified a form of natural selection known as *sexual selection.* The classic example of sexual selection is the peacock's tail. The function of the peacock's tail is not to promote the survival of the peacock; rather, the function is to promote the survival of the peacock's genes. Sexual selection arises once a particular genetic preference is established by the opposite sex -- in this case, the preference of the peahen for flashy tails. Even if one peahen is not particularly impressed by Las Vegas-style tails, it remains to the female's benefit to mate with the most colorful male if her offspring are more likely to be desired by other females who *are* fond of colorful tails.

Darwin himself suggested that music might have arisen due to sex selection in mating calls. Like the peacock's tail, the preferences of hominid women could create an escalating competition for ever more elaborate and beautiful melodies. The members of the all-male Vienna Philharmonic notwithstanding, there is nothing to indicate that one sex is more musical than the other, and so there is no evidence of the dimorphism symptomatic of sexual selection. Women may be impressed by men who serenade them outside their balcony windows, but it is questionable whether this says anything about evolutionary origins. After all, unlike female songbirds, female humans are perfectly capable of serenading men.

For the same reason, there is little to support the view that human music-making arose in a manner analogous to the songs of songbirds. In songbird species, only the male bird sings. That is, there is high sexual dimorphism for singing. Once again, in humans, there is no comparable sexual dimorphism.

Types of Evidence

In presenting a case for the evolutionary origins of music, we can consider five types of evidence:

> **Genetic evidence.** The best evidence of an evolutionary origin would be the identification of genes whose expression leads to the behavior in question. Unfortunately, it is rare for scientists to be able to link particular behaviors to specific genes. Although behavior-linked genes have been discovered in other animals (such as fruit flies), no behavior-linked gene has yet been conclusively established in humans. As in so many other areas, music has attracted a kind of folklore related to heritability. In some cultures, it is common for people to assume or believe that musical talent is partly inherited. More recently, work at the University of California, San Francisco by Baharloo *et al.* (1998) appears to suggest a genetic component for absolute pitch.

Biochemical evidence. Since genes are expressed in the form of proteins, we would expect to be able to identify proteins that influence musical behaviors. If we cannot find such proteins, then there is little likelihood that music has a genetic basis.

Neurological evidence. The existence of specialized brain structures is neither a sufficient nor a necessary condition for music to be an evolutionary adaptation. Nevertheless, if stable anatomical brain structures exist for music, then this is consistent with music arising from innate development rather than due solely to a generalized learning.

Ethological evidence. Are musical behaviors consistent with survival and the propagation of genes? In order for music to be an evolutionary adaptation, music-related behaviors must somehow increase the likelihood that the musical person's genes will be propagated.

Archaeological evidence. Since complex evolutionary adaptations arise over many thousands of generations, we must ask how widespread music is in biological history? If music originated in the past few thousand years, then it is highly unlikely to be an evolutionary adaptation. Evolution doesn't work that fast.

As noted, there is currently no evidence that links music to any gene. Let's consider the other areas of evidence in more detail.

Biochemical Evidence.

In 1980, Avram Goldstein published the results of an experiment where the effect of naloxone on musical pleasure was measured. Naloxone is an *opiate receptor antagonist* -- that is, it is a molecule that attaches to opiate receptors in the brain without activating them. Naloxone is a sort of prophylactic that covers the receptor and prevents it from being activated. In his experiment, Goldstein found that volunteer listeners who had been injected with naloxone reported significantly less musical pleasure than those who had received a saline injection.

Goldstein's experiment does not tell us *how* music evokes pleasure. But his experiment implies that, however music evokes pleasure, it ultimately causes the release of endorphins that stimulate the brain's opiate receptors. In short, musical pleasure appears to engage the same physiological mechanisms that are used by a wide variety of other pleasure-inducing behaviors.

Most activities that result in pleasure are somehow related to enhanced survival. As we already noted, there are some pleasure-inducing activities that do not appear to have any evolutionary or adaptive value. Is music adaptive pleasure-seeking or is it non-adaptive pleasure-seeking? That is, is music somehow akin to eating (an activity that increases survival)? Or is music like heroin use (an activity of no apparent survival value that simply exploits the biological pleasure-creating machinery designed for other purposes)? I would like to propose that this is one of the most fundamental questions that can be asked about music.

Further evidence related to biochemical concomitants of music was provided by Fukui (1996). Fukui measured the effect of music listening on testosterone production. Evidence supporting a social role can also be found in studies of physiological responses to music. Testosterone is an androgen -- a hormone normally associated with men, but also produced in smaller quantities in women. Testosterone levels are strongly correlated with aggression: high levels of testosterone tend to facilitate aggressive behaviors, similar to the "roid rage" commonly experienced by athletes who use commercial steroids. In addition, testosterone is thought to mediate libido. Low levels of testosterone are associated with lower sexual arousal. (Nelson, 1994; Sherwin, 1988; Wallen & Lovejoy, 1993)

Fukui (1996) carried out a study in which he measured testosterone levels from saliva samples collected from undergraduate-aged participants while they listened to their favorite music. Compared with a control group that

listened to no music, the testosterone levels dropped significantly. Moreover, Fukui found no sex-related differences: testosterone levels dropped by a similar proportion in both male and female listeners.

Both Fukui's experiment and Goldstein's experiment provide evidence that music modulates the production of specific proteins in the body. This doesn't prove much, but it does demonstrate that music is not a disembodied abstraction; music engages human physiology at one of the most basic levels. We will have more to say about Fukui's and Goldstein's experiments later.

Archaeological Evidence

Let's now move on to consider some of the archaeological facts. The archaeological record shows a continuous record of music-making in human settlements. Wherever you find evidence of human settlement, you find evidence of musical activities.

In 1995, paleontologist Ivan Turk discovered a bone flute while excavating an ancient burial mound in Divje Babe, Slovenia (Anon., 1997). Using electron spin dating, the age of this flute has been determined to lie between 43,000 and 82,000 years old. If the instrument had been made out of wood, it would long ago have disintegrated. So we are fortunate that someone took the time to fashion this particular instrument from the femur of the now extinct European bear.

Of course, finding this flute doesn't mean we've found the earliest musical instrument; this is just the earliest *found* instrument. It's logical to presume that wooden flutes were fashioned earlier than bone flutes. So it's not inconceivable that wooden flutes existed, say, 100,000 years or more ago.

As musical instruments go, flutes are rather complicated devices. If we look at contemporary hunter-gatherer societies, the most common instruments are rattles, shakers, and drums. For example, prior to the arrival of Europeans, by far the most common instruments in native American cultures were rattles and drums. The same pattern of preferred instruments is evident in African and in Polynesian cultures. If we assume that rattles and drums typically pre-dated the use of flutes, then the ancient music-makers of Slovenia might well have been creating instrumental music somewhat earlier than 100,000 years ago.

But what sort of music-making might have existed prior to the fashioning of musical instruments? It's not unreasonable to assume that singing preceded the making of musical instruments by some length of time. If we suppose that singing predated instrument making by 50% of the intervening time, then music-making might have existed 150,000 years ago -- roughly twice the age of the older estimate for the Divje Babe flute. Even this figure might be a conservative estimate, and the actual origin of music might be twice as old, say around 250,000 years ago.

On the other hand, the Divje Babe flute might truly be an early specimen, and singing might have developed about the same time. Using the most recent estimate for the Divje Babe flute would therefore place the origins of music-making about 50,000 years ago.

In summary, the archaeological record implies that music-making likely originated between 50,000 years ago and a quarter of a million years ago. Although Wurlitzer organs, American Bandstand, and MTV are relatively recent phenomena, music-making in general is really quite old.

The evidence pointing to the great antiquity of music satisfies the most basic requirement for any evolutionary argument. Evolution proceeds at a very slow pace so nearly all adaptations must be extremely old. Music-making satisfies this condition, though we must be careful not to assume that the music of the Pleistocene period bears much resemblance to Brahms or Twisted Sister.

Anthropological Evidence

Turning to contemporary anthropology, we can ask, "What does the plethora of existing human cultures tell us about music?" Without taking time to review the evidence, there is one overwhelming conclusion from the modern anthropological record. There is no human culture known in modern times that did not, or does not, engage in recognizably musical activities.

Not only is music-making very old, it is ubiquitous; it is found wherever humans are found. Moreover, I neglected earlier to mention one important fact about the bone flute at Divje Babe: the flute was found in a Neanderthal burial site. The Divje Babe flute isn't even a human artifact. In short, it may be the case that music-making is not just ubiquitous among *homo sapiens*; music-making may possibly be characteristic of the entire genus *homo*.

The evidence pointing to the ubiquity of music satisfies another important basic requirement for any evolutionary argument. Relatively few adaptations are *not* found throughout the entire population of the affected species. For example, if eyelashes confer an evolutionary advantage, then just about everyone should have eyelashes. There are some exceptions to this principle, some of which are very important. For example, humans divide into female and male versions, so there are some genes that are not shared by everyone. Another more subtle example is the gene that codes for sickle-cell -- a gene that protects against malaria, but can also cause anemia.

Ethological Evidence

Ethology is the study of animal behavior; this includes the study of human behavior. When studying a particular animal, ethologists often begin by making an inventory of observed behaviors. What does the animal do, and how often does it do it? Activities that require a great deal of time and large expenditures of energy are understandably considered important. Ethologists assume that behaviors are likely to be optimized. Even behaviors that seem unimportant (such as infant play, or sleeping) often have a serious or critical purpose.

Primates, for example, spend an extraordinary amount of time grooming each other. Ethologists feel obliged to formulate theories that account for the various proportions of resources dedicated by an animal to different activities.

Let's apply the ethological approach to the behaviors we call musical. For the purposes of illustration, let's consider two case descriptions. The first case is that of the Mekranoti Indians of the Brazilian Amazon. And the second case is that of contemporary U.S. society.

The Mekranoti Indians

The Mekranoti Indians are hunter-gatherers who live in the Amazon rainforest of Brazil. In Mekranoti culture, singing plays a prominent role in daily life. For several months of the year, every morning and evening the women lay banana leaves on the ground where they sit and sing for between one and two hours. The men sing every night starting typically around 4:30 in the morning, but sometimes as early as 1:30 AM. The men sing for roughly two hours each night, and often they will also sing for a half hour or so before sunset.

When singing, the Mekranoti men hold their arms in a sort of cradling position and swing their arms vigorously. The men endeavor to sing in their deepest bass voices, and heavily accent the first beats of a pervasive quadruple meter with glottal stops that make their stomachs convulse in rhythm. Anthropologist Dennis Werner (1984) describes their singing as a "masculine roar." When gathering in the middle of the night, the men are obviously sleepy, and some men will linger in their lean-tos well after the singing has started. These malingerers are often taunted with shouted insults.

Werner reports that "Hounding the men still in their lean-tos [is] one of the favorite diversions of the singers. `Get out of bed! The Kreen Akrore Indians have already attacked and you're still sleeping,' they [shout] as loudly

as they [can]. ... Sometimes the harassment [is] personal as the singers [yell] out insults at specific men who rarely [show] up." (pp.245-247)

What is extraordinary about the Mekranoti singing is the amount of time involved -- roughly two hours per day. (Remember, this is a subsistence hunter-gatherer society.) For the evolutionary ethologist, the important question arising from the Mekranoti Indians is why music-making would attract so much of the tribe's resources. We'll return to this question later.

Modern U.S.

By way of comparison, consider now the prevalence of music in a modern industrialized society like the United States. For the ethologist looking at modern human behaviors, a crude though ready index of the amount of resources we dedicate to a particular activity can be found by measuring economic activity.

There is a widespread misconception that the foremost export sector in the U.S. economy is "high technology." In fact, the preeminent export sector in the U.S. economy is entertainment. Of the various component areas -- films, sports, television, toys, and games -- it is music that ranks foremost.

How big is the music industry? The music industry is bigger than the pharmaceutical industry. People spend more money on music than on prescription drugs. We purchase recordings, go to concerts, buy sheet music, take our children to music lessons, listen to commercial radio, watch film accompanied by music, and encounter Muzak in the local shopping mall. The most active `concert venues' in the world are freeways: a major preoccupation for millions of drivers is listening to music.

Of course financial measures are crude indicators of behavioral significance. The ethological point is simple. In both a hunter-gatherer society and a modern industrial society, we find humans dedicating a notable proportion of resources to music-making and listening. Music may not be more important than sex; but it is arguably more expensive, and it is certainly more time-consuming.

In order to put these behaviors in perspective, suppose you were a Martian anthropologist visiting earth. There are many aspects of human behavior that would have recognizable value. You would see people engaged in growing and preparing food, in raising and educating children, people involved in transportation, health, and governance. But even if Martian anthropologists had ears, I suspect they'd be stumped by music.

If you're still not convinced that music attracts a peculiarly excessive proportion of human resources, consider another comparison. Think of how important food is to human well-being; of how tasty and enjoyable food is and can be. Now how many universities have departments of cuisine or nutrition? Or departments of food sciences, or even departments of home economics? Now consider how many university have departments of music. Why would music figure more prominently than food? To a visiting tourist from Mars, *music sticks out*; it is a remarkable and bizarre activity that earthlings do.

Of course, we must be careful in drawing any conclusions about adaptations based on observations of modern behaviors. If music-making is an adaptive behavior, then it must have arisen long ago in the environment of evolutionary adaptedness -- namely the Pleistocene period when the vast majority of human evolution occurred.

Ethology and Evolution

Just because an animal spends a lot of time on certain activities doesn't mean that the activity represents an evolutionary adaptation. Ethologists must connect the behavior to an explicit evolutionary account. That is, there must exist a plausible explanation of how the behavior would be adaptive.

Before considering such a theory for music, let's examine a non-musical example -- an example that has a richer theoretical literature about its origins. Specifically, let's consider some of the evolutionary arguments that have

been advanced to account for the origins of language.

On the Evolutionary Origin of Language

As in the case of music, views concerning the origins of language are necessarily speculative. Nevertheless, we can learn a great deal by considering some of the theories that have been advanced concerning its origin. Until recently, the principal view of language was that it facilitated complex collaborative activities such as coordinating actions during hunting. This account seems unlikely, first, because talking is bad idea when tracking prey, and secondly, because men as a group display inferior language skills compared with women.

A number of anthropological psychologists have suggested that language (and even music) evolved as surrogates for social bonding.

The Grooming and Gossip Theory of Language Origins

The most empirically grounded of the recent theories of language origins is what might be called the "grooming and gossip hypothesis." Its principal advocate is Robin Dunbar (1997). The theory proposes the following logic.

Animals often live in groups for mutual protection against predators. In general, larger groups are more effective in detecting and warding-off predators than smaller groups. But there are costs associated with maintaining a large group. One cost is that feeding must be much more intensive in a given area and so a larger group must travel greater distances in search of food. A second cost is that as group size increases, threats are more likely to arise from internal conflict within the group rather than from external predators. That is, there is a point where group size effectively minimizes predation, but at the cost of threats from members of the group itself. Nowhere is this more evident than in primates. As a consequence of internal threats, animals within the group begin to form alliances with one another. These alliances reduce the likelihood of conflict due to the threat of group retaliation.

In primates, the principal means by which alliances are formed and bonds maintained is through grooming. Grooming accounts for between 10% and 20% of an individual's daytime activities.

There is good evidence to suggest that the principal purpose of grooming is to form alliances between individuals. First, grooming partners are much more likely to come to the defense of one-another when threatened by another member of the group. Even more important evidence comes from relating the amount of time spent grooming to the size of the group. Different primate species have different typical group sizes. Gorillas, macaques, chimpanzees, bonobos, and so on each tend to form groups that have different average sizes. Primatologists have measured the different amounts of time each species engages in grooming.

A major discovery has been that there is a consistent relationship between group size and the amount of time spent grooming. As the group size increases, the average grooming time also increases. This is an important finding: there is no reason to suppose that animals in larger groups tend to get more dirty than animals in smaller groups, so the increase in grooming is unlikely to be related to cleanliness.

Primatologists widely agree that the increase in grooming time for larger groups arises from the need to form more extensive networks of alliances. In a large group, an individual fares better by having a wider circle of friends, and the way to build primate friendships is through mutual grooming.

Of course alliances can be broken or betrayed. An animal who has been attacked by another animal may well expect a grooming partner to come to their defense. But there are always those individuals who may benefit from your willingness to defend them, but who will not reciprocate by coming to your defense. This is the so-called "free-rider" problem: cunning animals might well exploit those who are foolish enough to groom them.

The free-rider problem means that individual primates ought to be sensitive to the possibility of defection by a grooming partner. Individuals will look for clues about the reliability of those they consider their friends. Indeed, primatologists have described circumstances where a grooming alliance is abandoned by an individual who has witnessed the failure of their partner to come to the defense of a third grooming partner. Untrustworthy animals are not popular grooming partners, and a reputation for reciprocal altruism is important.

In this respect, humans are no different from other primates. As Cosmides and Tooby (1992) have famously shown, human reasoning follows patterns, not of abstract logic, but are optimized for social contracts. Humans have deep-seated notions of justice that follow from betrayals of social alliances: if you collaborate, you deserve to be helped; if you defect, you ought to pay the price.

In the case of humans, the common `group size' has been estimated at roughly 150 people. This is approximately the size of most rural villages in the world. This means that human groups are especially large when compared with other primates. As Dunbar (1997) has pointed out, "If modern humans tried to use grooming as the sole means of reinforcing their social bonds, as other primates do, then the equation for monkeys and apes suggests we would have to devote around 40 per cent of our day in mutual mauling." (p.78).

Dunbar has suggested that language evolved as an alternative to physical grooming. In effect, physical grooming was replaced by "vocal grooming" whose purpose remains the formation and maintenance of friendships or alliances. Such "vocal grooming" has two distinct advantages over physical grooming. First, we can talk to several people simultaneously. This increases the number of people we can bond with at the same time. Second, we can exchange information about people who are physically absent -- that is, we can gossip. Unlike other primates, this means that we can learn about the behavior of others without being limited to direct observation.

Incidentally, Dunbar's theory does not preclude other *uses* for language. Clearly, language is advantageous in a number of ways. Dunbar's theory simply attempts to account for how language got started in the first place -- it is not necessarily a theory of how language might be adaptive for modern humans.

Nevertheless, Dunbar and his colleagues have conducted a number of studies that illustrate the continuing human penchant for gossip. Even in formal business interactions, only roughly 1/4 of the time is spent negotiating or discussing technical details. The majority of time in business interactions is spent relaying personal information, discussing colleagues, and gossiping about the intentions, betrayals, supports, and reliability of other people -- or establishing our own credibility and worthiness of character.

When did language arise in humans? Estimates vary from as recently as 50,000 years ago to 500,000 years ago. None of the evidence is direct. Archaeologists point to the so-called Upper Paleolithic Revolution, a period when stone artifacts and tools show a marked improvement in quality and range. At this point (50,000 years ago) tools include buttons, needles, awls, and other refined inventions.

Direct evidence for language (such as writing) is only found within the last 10,000 years. In fact, the archaeological evidence for the antiquity of music is stronger than the archaeological evidence for the antiquity of language -- although that doesn't mean that music necessarily preceded the emergence of language.

Note that even language has significant limitations for multiple concurrent social interaction. Dunbar (1997) has noted that "there appears to be a decisive upper limit of about four on the number of individuals who can be involved in a conversation." (p.121). When a fifth or sixth person joins a conversation there is a marked tendency for the group to subdivide into two or more concurrent conversations. It is only in hierarchical situations (such as in a formal lecture) where a single conversation can be maintained in a larger group.

All of this suggests that language is most useful in close interpersonal interactions, such as grooming, gossiping, courting, and conspiring. Note, however, that there are other activities that are of value to members of a social group that involve the entire group (or at least large segments) rather than groups of twos or threes. Chief among these group activities is defense. When under threat, uniform group action is indeed a mighty force, much more powerful than smaller groups of twos and threes.

Music and Social Bonding

At this point, we might speculate how music might fit into this account. Let's assume, for the moment, that the hypothesis that language evolved as a surrogate for physical grooming is true, and that language thereby allowed humans to live in larger groups with their attendant complex social relations. We could certainly conceive of a similar function for music. In some ways, music provides several advantages over language. Singing is much louder than speaking, so singing may facilitate group interactions involving more than the four individuals posited as the upper limit for conversation. Although music may not be as effective as language in informing us of the deceptions of others, it does fit within the rubric of surrogate grooming. Recall that in primates, the function of grooming is to provide social bonding opportunities -- not ways of learning about the machinations of individuals who are absent. Therefore, in some ways, music provides a better parallel to physical grooming than is the case for language.

This view of the possible origins for music was essentially proposed by Juan Roederer (1984):

> "... the role of music in superstitious or sexual rites, religion, ideological proselytism, and military arousal clearly demonstrates the value of music as a means of establishing behavioral coherency in masses of people. In the distant past this could indeed have had an important survival value, as an increasingly complex human environment demanded *coherent,* collective actions on the part of groups of human society." (p.356)

In light of later work by primatologists such as Dunbar, there appears to be merit in Roederer's hypothesis. Music might have originated as an adaptation for social bonding -- more particularly, as a way of synchronizing the mood of many individuals in a larger group. That is, music helps to prepare the group to act in unison.

Perhaps a helpful image is to imagine the cackling of geese prior to them taking off. How is it that individual geese manage to synchronize their actions so that the entire flock takes flight more-or-less simultaneously? For anyone who has watched geese take-off, there is a clear increase in the volume of cackling ... more and more geese start honking. The general hubbub of honking geese is apt to raise the arousal levels of all geese in the vicinity. This increased arousal (which includes increased heart-rate) would prepare the geese for a significant collective expenditure of energy.

Music and Social Bonding -- Further Evidence

It is this theory of music and social bonding which I believe holds the greatest promise as a plausible evolutionary origin for music. For the remainder of this lecture, I would like to review further phenomena that provide support for this hypothesis. The evidence is going to come from the following five sources:

1. Various mental disorders imply a strong link between sociability and musicality.
2. Child development implies a social role for music.
3. Brain structures related to music are linked to social and interpersonal functions.
4. The most popular musical works imply social functioning.
5. Music modifies hormone productions in groups of people.

Complementary Disorders: Williams Syndrome and Asperger Autism

Consider two mental disorders: Williams Syndrome and Asperger-type Autism. The principal feature of Williams syndrome is mental retardation. Williams syndrome is somewhat unique in that suffers display three additional characteristics. One characteristic is high verbal abilities. Individuals suffering from Williams syndrome take a great interest in words. Their speech is fluent, and peppered with a remarkably sophisticated vocabulary. In fact, when first encountering someone with Williams syndrome the language fluency tends to mask the mental handicap.

In addition to high verbal abilities, Williams syndrome individuals also exhibit high sociability. They are gregarious and sociable. Coupled with the high verbal abilities, this makes Williams syndrome children a delight to work with. Finally, Williams syndrome children exhibit high musicality.

Daniel Levitin and Ursula Bellugi (1997) have described the musical activities of Williams syndrome children at a summer camp in New York state. The children are remarkable. The entire camp is alive with music, string quartets, trios, woodwind groups, and so on. They are `crazy' about music, and relish the social environment of other children with the same social, language, and musical enthusiasms.

Now consider the case of Asperger-type Autism. Autism is characterized by a strong aversion to social interaction. Although most autism is associated with reduced mental functioning, mental retardation is not always evident. There are autistic individuals with normal and above average intelligence as well. Autism is related to an emotional deficit -- notably the failure to develop the so-called *secondary* or *social* emotions -- including shame, pride, guilt, love and empathy. For normal children, these secondary emotions typically appear by about the age of four.

Temple Grandin is a high functioning Asperger-type autistic who has become well-known through her writings about her own condition. Concerning love, Grandin talks about her confusion in high school when reading Shakespeare's *Romeo and Juliet.* I never figured-out what it was all about, says Grandin. In a trip through the Rocky Mountains with Oliver Sacks, Grandin remarked "The mountains are pretty, ... but they don't give me a special feeling, the feeling you seem to enjoy." "You get such joy out of the sunset," she said. "I wish I did, too. I know it's beautiful, but I don't `get' it." (Sacks, p.124). Grandin's experience of music is similar. Although Grandin has perfect pitch, and what she describes as a tenacious and accurate auditory memory, she finds music leaves her cold. She finds the sounds "pretty," but in general, she just doesn't "get" it (p.122). All the fuss about music leaves her mystified.

Grandin's own explanation is that not all of the `emotional circuits' are connected. Sack's interprets the phenomenon as follows: "An autistic person can have violent passions, intensely charged fixations and fascinations, or, like Temple [Grandin], an almost overwhelming tenderness and concern in certain areas. In autism, it is not affect in general that is faulty but affect in relation to complex human experiences, social ones predominantly, but perhaps allied ones -- esthetic, poetic, symbolic, etc. No one, indeed, brings this out more clearly than Temple herself. ... She feels that there is something mechanical about her mind, and she often compares it to a computer ... She feels that there are usually genetic determinants in autism; she suspects that her own father, who was remote, pedantic, and socially inept, had Asperger's -- or, at least, autistic traits -- and that such traits occur with significant frequency in the parents and grandparents of autistic children." (p.123).

The contrast between Asperger-type Autism and Williams Syndrome is striking. On the one hand we have a group of people whose symptoms include high sociability linked with high musicality. On the other hand we have a group of people whose symptoms include low sociability OFTEN linked with low musicality. Together, these mental conditions are consistent with a relationship between sociability and musicality -- and this link is the principal assumption of a group-oriented evolutionary account.

Music and Social Function

Suppose we asked the following question: What is the most successful piece of music in modern history? Of course the answer to this question depends on how we define success -- and this is far from clear, as esthetic philosophers have shown. Nevertheless, let's use a straightforward criterion: lets assume that the most successful musical work is the one which is most performed and most heard. Using this criterion, you might be surprised by the answer. The most successful musical work was composed by Mildred and Patti Hill in 1893, and revised in the 1930s (Fuld, 1995). The piece in question is, of course, *Happy Birthday*. *Happy Birthday* has been translated into innumerable languages and is performed on the order of a million times a day. It remained under copyright protection until the middle of the century. For many people, the singing of *Happy Birthday* is the only time they sing in public. For other people, the singing of *Happy Birthday* constitutes the only time they sing.

In some ways, *Happy Birthday* is the quintessential feminist work. Its composers remain unknown and uncelebrated; the work was created by the collaboration of two women rather than as an egotistic expression of one man. It is a thoroughly domestic work; *Happy Birthday* is performed in the kitchen or lunch room rather than in the concert hall. No other musical work has evoked so much spontaneous music-making. The work is domestic, amateur, and relationally oriented. Despite its extraordinary success, it remains undervalued as a musical creation.

Happy Birthday plays a role in our evolutionary story because I suspect that for the vast majority of human history, music-making was of this ilk. In Western culture, it is surely the camp songs sung by Girl Scouts or the songs sung by British soccer hooligans that come closest to what might be imagined in Pleistocene homo sapiens. In all of these cases, the music serves an obvious social role and is a critical moment in defining a sense of identity and common purpose.

In light of our evolutionary hypothesis, let's return and reconsider the singing of the Mekranoti Indians. Recall some of the characteristic features -- especially the singing done by the men: the men's singing is done late at night and in the early morning, and their singing is associated with a high degree of machismo. Like most native societies, the greatest danger facing the Mekranoti Indians is the possibility of being attacked by another human group. The best strategic time to attack is in the very early morning while people are asleep. Recall the insult shouted at men who continued to sleep in their lean-tos: "Get out of bed! The Kreen Akrore Indians have already attacked and you're still sleeping."

The implication is obvious. It appears that the nightly singing by the men constitutes a defensive vigil. The singing maintains arousal levels and keeps the men awake.

Of course music-making is also associated with stirring a war-party. North American Indians famously sang and danced prior to initiating an attack on another tribe. One might suppose that engaging in an activity that publicly announces a hostile intention would be counter-productive. War dances might possible warn an enemy of an impending attack. However, the music-making seems to serve a more important role: that of raising arousal and synchronizing individual moods to serve the larger goal of the group.

Social Bonding and Hormones

Apart from arousing individuals, music can also pacify. Recall Fukui's experiment showing that listening to music can reduce testosterone levels. Fukui himself was quick to point out the possible social and evolutionary significance of this finding. In human social groups, lower levels of testosterone are likely to result in less aggression, less conflict, less sexual confrontation or sexual competition, and consequently more group cohesiveness. Where men commonly suffer from testosterone "poisoning," music truly hath charms to soothe the savage breast.

A problem with Fukui's experiment is that he didn't manipulate the type of music heard by his listeners. Listeners simply listened to their *favorite* music. Depending on his sample of listeners, we might expect whole genres of music that were not represented. We might suppose, for example, that heavy metal, hard rock, or thrash music might well have increased testosterone levels rather than decreasing the levels. Further research is necessary to document the specific hormonal changes associated with different types of musical experiences. However, Fukui's work at least shows that music can have marked effects on hormone levels -- specifically, hormones that relate especially strongly to sociability.

Oxytocin and the Biology of Social Bonding

An important question to ask is how precisely music might bring about social bonding. Neurophysiologist Walter Freeman (1995) has proposed a pertinent theory related to the hormone *oxytocin*.

Oxytocin is most commonly associated with the "let-down" response in new mothers -- that is, the response that enables the flow of breast milk following child-birth. The presence of oxytocin also has dramatic effects on the brain. For example, when a ewe gives birth to a lamb, the olfactory bulb in the ewe's brain is bathed in oxytocin. Following the birth of the new lamb, a ewe will imprint on the smell of the new lamb, but will subsequently fail to recognize the smell of her former offspring. The result is that the ewe will suckle only the new-born lamb.

Neurophysiological research has shown that oxytocin acts as a sort of "eraser" that wipes away previous memories and simultaneously facilitates the storage of new memories. When linked with significant life events, oxytocin is the cement that binds new memories. The amnesic properties of oxytocin are evident in all kinds of learning episodes. However, their strongest effects occur during major limbic activations such as those resulting from trauma or from ecstasy. Pavlov discovered this phenomenon when serious spring flooding affected his lab and nearly drowned his caged dogs. Following their rescue it was discovered that the dogs had to be re-trained from scratch (Pavlov, 1955).

In his book, *Societies of Brains,* Freeman chronicles a number of circumstances where oxytocin release occurs, and the effects of these releases on neural organization. As we have noted, oxytocin releases are associated with trauma and ecstasy. In addition to child-birth, oxytocin is released in males and females following sexual orgasm. Freeman also suggests that oxytocin is released during trance and while listening to music.

In many cases, the presence of oxytocin is correlated with human and animal bonding circumstances. For example, in the case of sexual orgasm, oxytocin may significantly facilitate pair-bonding in the same way that oxytocin following child-birth facilitates mother-child bonding. Freeman's suggestion that music causes oxytocin to be released has important repercussions for instances of peer-group bonding and social identity. If Freeman is correct, there would be good neurophysiological reasons for lovers to enjoy music while courting, for union members to sing while on the picket line, for religious groups to engage in collective music-making, for colleges to promote *alma mater* songs, and for warriors to sing and dance prior to fighting.

Mood Regulation

Thayer and his colleagues have carried out a number of studies concerning how people regulate their moods. One study attempted to determine what people do to try to get out of a bad mood. Of 29 categories of activities, the foremost activity was calling or talking to a friend. The second most frequently reported activity was trying to think positive thoughts -- to give oneself a sort of "pep talk." The third most frequently reported activity -- ahead of a wide variety of behaviors -- was listening to music. Forty-seven percent of respondents reported that they used music to temper or eliminate a bad mood.

Thayer *et al.* carried out a similar study to determine what people do to raise their alertness or energy level. Listening to music was reported by 41 percent of respondents, following activities such as sleeping, taking a shower, getting some fresh air, and drinking coffee. Finally, in a third study investigating what people do to reduce nervousness, tension or anxiety, listening to music ranked third at 53 percent, following after only calling or talking to someone, and trying to calm down by thinking about a situation.

There are two points to highlight from these studies. The first is that the foremost category of behavior for mood regulation is *being with or conversing with a friend.* That is to say, our first tendency is to seek mood regulation through *social interaction.* Moods are contagious, and we rely to some extent on each other to to modulate, reinforce or temper our moods. Although we know that moods are highly influenced by the individual's physiological state -- notably through food, exercise, rest, etc. -- behaviors such as eating, exercise and rest are less frequently used for mood regulation than music.

The second point to highlight is the obvious point that music appears to figure prominently as a method for mood regulation. Although in contemporary society music tends to be experienced in a personalized or individualized listening context, we already know that this context is historically unprecedented. Most music-making in hunter-gatherer societies occurs in a social or group context. Until the invention of the phonograph,

the vast majority of music in Western culture was also experienced in social or group contexts. In short, music is not out-of-place in the list of socialized behaviors used for mood regulation.

Conclusion

By way of conclusion, first let me again reiterate that I don't think the evidence in support of music as an evolutionary adaptation is strong. The purpose of this lecture has been to show that there are no obvious or fatal impediments that rule-out a possible evolutionary origin.

We might summarize the basic evidence as follows:

1. Complex evolutionary adaptations arise only over many millennia. Accordingly, in order for a behavior to be adaptive, it must be very old. As we have seen, music-making does indeed conform to the criterion of great antiquity.
2. Evolution proceeds only by changes to a species' genome. Evolution influences genes, and genes are expressed in the form of proteins, so any purported adaptation must have biochemical concomitants. As we have seen, musical experience clearly influences and is modified by natural biochemical substances in the body. Music evokes pleasure via the same ultimate pathway as for other forms of behavior, and music regulates the production of testosterone and (possibly) oxytocin. These facts in no way prove that music is an adaptation, but they satisfy a basic biochemical requirement.
3. Behavioral specializations are often expected to be associated with specific anatomical or functional brain structures. Lesions and other neurological assaults can leave an individual with impaired musical functioning. There are double-dissociations between various amusias and virtually every other kind of functional mental loss. This does not prove that music is not acquired by general learning, but the neurological evidence is at least consistent with the possibility that there are specialized music-related brain structures.
4. In order for a behavior to be adaptive, the behavior itself must enhance the propagation of the individual's genes. As we have seen, musical behaviors are consistent with mood modification and group mood synchronization -- and these synchronous states are at times clearly associated with situations where group efforts are adaptive -- such as in the case of defense against other human groups. In addition, high musical involvement is not associated with dereliction or poor survival (such as the case for alcohol); this raises problems for the view that music is a form of non-adaptive pleasure-seeking.

The evidence we have for mood regulation and synchronization is suggestive:

1. We have noted contrasting disorders in Williams syndrome and Asperger-type Autism. In one case, we see a group of individuals who are highly sociable and also highly musical. In the other case, we see /a group of/ SOME individuals who display extremely low sociability and also low musical understanding or affinity.
2. Although we didn't review this literature, the emergence of the secondary or socialized emotions in child development is strongly associated with musical empathy, understanding and sophistication. The pertinent research on child development implies a social role for music.
3. We noted that the most popular musical works often imply some sort of social function. *Happy Birthday* is only one example. Group identity is often expressed through folk songs, Girl Scouts camp songs, sports, war dances, and so on.
4. Although we didn't review the literature, it is also known that the emergence of musical tastes relates to post-pubescent socializing and group identity.
5. And finally, we discussed how music modifies hormone productions in groups of people.

As noted at the beginning of this essay, there is a long history of abuse of genetic claims serving ulterior and often nefarious motives. Even if we assume that musicality has some adaptive function, the repercussions for modern music-making and modern musical enjoyment are likely to be minimal.

Music is now deeply embedded in a cultural/historical context where human musical memories span centuries and the fashion cycle is a significant engine of change. Music is now part of a Lamarkian system where acquired characteristics are transmitted in Dawkinsean "meme-pool" rather than in Mendelian "gene-pool". Like language, the details of musical culture and tastes are largely a product of enculturation.

Nevertheless, it remains worthwhile to attempt to understand where music comes from, and why it has achieved such a ubiquitous presence in human lives. Evolutionary theorizing about music may well remain in the realm of `Just-So' stories. But there is always the possibility of a testable hypothesis emerging, and if so, we'll all wait with interest to see the results.

Acknowledgments:

I would like to extend my thanks to Dr. Kristin Precoda for drawing my attention to the work of Werner regarding the Mekranoti Indians, and to Dr. David Wessel for drawing my attention to Freeman's work concerning oxytocin. This lecture was originally presented at the Department of Music, University of California, Santa Barbara, March 6th, 1998. A shortened version was presented at the Society for Music Perception and Cognition Conference, Evanston, Illinois. August 16, 1999. An edited version of this article can be found in: D. Huron (2001), "Is music an evolutionary adaptation?" *Annals of the New York Academy of Sciences*, Vol. 930, pp. 43-61.

References:

Anon. (1997).
Neanderthal notes: Did ancient humans play modern scales? *Scientific American,* Vol. 277, No. 3 (1997, Sept.) pp.28-30.

Aiello, L., & Dunbar, R.I.M. (1993).
Neocortex size, group size and the evolution of language. *Current Anthropology,* Vol. 34, pp. 184-193.

Baharloo, S., Johnston, P.A., Service, S.K., Gitschier, J. & Freimer, N.B. (1998).
Absolute pitch: An approach for identification of genetic and nongenetic components. *American Journal of Human Genetics,* Vol. 62, pp. 224-231.

Barkow, J., Cosmides, L. & Tooby, J. (Eds.) (1992).
The Adapted Mind: Evolutionary Psychology and the Generation of Culture. Oxford: Oxford University Press.

Baron-Cohen, S. (Ed.). (1997).
The Maladapted Mind: Classic Readings in Evolutionary Psychopathology. East Sussex, UK: Psychology Press.

Blair, R.J.R. (1995).
A cognitive developmental approach to morality: investigating the psychopath. *Cognition: International Journal of Cognitive Science,* Vol. 57, pp. 1-29.

Brown, D. (1991).
Human Universals. New York: McGraw-Hill.

Buss, D. (1994).
The Evolution of Desire: Strategies of Human Mating. New York: Basic Books.

Cacioppo, J.T., Petty, R.P. Losch, M.E., & Kim. H.S. (1986).
Electromyographic activity over facial muscle regions can differentiate the valence and intensity of affective reactions. *Journal of Personality and Social Psychology,* Vol. 50, pp. 260-268.

Darwin, Charles. (1872).
The Expression of Emotion in Man and Animals. London: Murray.

Dawkins, R. (1976).
The Selfish Gene. Oxford: Oxford University Press.

Dawkins, R. (1982).
The Extended Phenotype. San Francisco: W.H. Freeman.

Dawkins, R. (1986).
The Blind Watchmaker. New York: W.W. Norton.

Dawkins, R. (1995).
River Out of Eden. New York: Basic Books.

Dimberg, Ulf. (1987).
Facial reactions and autonomic activity to auditory stimuli with high and low intensity. *Psychophysiology,* Vol. 24, p. 586.

Dimberg, Ulf. (1989).
Perceived unpleasantness and facial reactions to auditory stimuli. Uppsala, Sweden: Uppsala Psychological Reports, No. 414.

Dissanayake, Ellen. (1988).
What Is Art For? Seattle: University of Washington Press.

Dobzhansky, Th. (1973).
Genetic Diversity and Human Equality. New York: Basic Books.

Dunbar, R. (1997).
Grooming, Gossip and the Evolution of Language. New York: Faber & Faber.

Durham, W. (1991).
Coevolution: Genes, Culture, and Human Diversity. Stanford, CA: Stanford University Press.

Eaves, L., Eysenck, H., & Martin, N. (1989).
Genes, Culture and Personality. London: Academic Press.

Ekman, P. (1973).
Cross-cultural studies of facial expressions. In P. Ekman (editor), *Darwin and facial expression.* New York: Academic Press, pp. 169-222.

Eysenck, H. (1997).
The future of psychology. In: R.L. Solso (ed.), *Mind and Brain Sciences in the 21st Century.* Cambridge, MA: MIT Press, pp.271-301.

Freeman, W. J. (1995).
Societies of Brains: A Study in the Neuroscience of Love and Hate. Hillsdale, NJ: Lawrence Erlbaum Associates.

Freeman, W. J. & Grajski, K. A. (1987)
Relation of olfactory EEG to behavior: Factor analysis. *Behavioral Neuroscience,* Vol. 101, pp. 766-777.

Fuld, J.J. (1995).
The Book of World-famous Music; Classical, Popular and Folk. New York: Dover, 4th edition.

Gaulin, S.J.C. & Robbins, C.J. (1991).
Trivers-Willard effect in contemporary North American society." *American Journal of Physical Anthropology,* Vol. 85, pp. 61-69.

Goldstein, A. (1980).
Thrills in response to music and other stimuli. *Physiological Psychology,* Vol. 8, No. 1, pp. 126-129.

Gould, S.J., & Lewontin, R.C. (1979).
The spandrels of San Marco and the Panglossian program: A critique of the adaptationist programme. *Proceedings of the Royal Society of London,* Vol. 250, pp. 281-288.

Hatfield, E., Cacioppo, J., & Rapson, R. (1994).
Emotional Contagion; Studies in Emotion and Social Interaction. Cambridge: Cambridge University Press.

Levitin, D.L., & Bellugi, U. (1997).
Musical abilities in individuals with Williams Syndrome. Paper presented at the 1997 Society for Music Perception and Cognition, Massachusetts Institute of Technology, Cambridge, Mass.

Lewontin, R.C. (1991).
Biology as Ideology; The Doctrine of DNA. Concord, Ontario: House of Anansi Press.

Miller, W.I. (1997).
The Anatomy of Disgust. Cambridge, MA: Harvard University Press.

Nesse, R
What good is feeling bad? *The Sciences.*

Pavlov, I. P. (1955).
Selected Works. Moscow: Foreign Languages Publishing House. Translated by S. Belsky; edited by J. Gibbons.

Pert, C.B. (1997).
Molecules of Emotion. New York: Scribner.

Pinker, S. (1994).
The Language Instinct. New York: Morrow.

Popper, K. (1935/1959).
Logik der Forschung. Vienna, 1935. Translated as *The Logic of Scientific Discovery.* New York: Basic Books.

Rozin, P. & Fallon, A. (1987).
A perspective on disgust. *Psychological Review,* Vol. 94, pp. 23-41.

Rozin, P., Haidt, J. & McCauley, C. (1993).
Disgust. In M. Lewis & J. Haviland (editors), *Handbook of Emotions.* New York: Guilford, pp. 575-594.

Ridley, M. (1994).
The Red Queen: Sex and the Evolution of Human Nature. New York: Macmillan.

Roederer, J. (1979).
Introduction to the Physics and Psychophysics of Music. New York: Springer Verlag.

Roederer, J. (1982).
Physical and neuropsychological foundations of music. In M. Clynes (Ed.), *Music, Mind and Brain.* New York and London: Plenum Press.

Roederer, J. (1984).
The search for a survival value of music. *Music Perception,* Vol. 1, No. 3, pp. 350-356.

Shepard, R.N. (1987).
Evolution of a mesh between principles of the mind and regularities of the world. In J. Dupr\o'e\(aa' (Ed.), *The Latest on the Best: Essays on Evolution and Optimality.* Cambridge, MA: MIT Press/Bradford Books; pp. 251-275.

Shepard, R.N. (1992).
The perceptual organization of colors: An adaptation to regularities of the terrestrial world? In: Barkow, Cosmides & Tooby (Eds.), *The Adapted Mind: Evolutionary Psychology and the Generation of Culture.* Oxford: Oxford University Press; pp. 495-532.

Symons, D. (1992).
On the use and misuse of Darwinism in the study of human behavior. In: Barkow, Cosmides & Tooby (Eds.), *The Adapted Mind: Evolutionary Psychology and the Generation of Culture.* Oxford: Oxford University Press; pp. 137-159.

Tooby, J., & Cosmides, L. (1992).
The psychological foundations of culture. In Barkow, Cosmides & Tooby (Eds.), *The Adapted Mind: Evolutionary Psychology and the Generation of Culture.* Oxford: Oxford University Press, pp. 19-136.

Trivers, R.L. & Willard, D.E. (1993).
Natural selection of parental ability to vary the sex ratio of offspring. *Science,* Vol. 179, pp. 90-91.

Wallin, N.L. (1991).
Biomusicology; Neurophysiological, Neuropsychological, an Evolutionary Perspectives on the Origins and Purposes of Music. Stuyvesant, NY: Pendragon Press.

Werner, D. (1984).
Amazon Journey; An anthropologist's year among Brazil's Mekranoti Indians. New York: Simon and Schuster.

Williams, L.
The Dancing Chimpanzee: A Study of the Origin of Music in Relation to the Vocalising and Rhythmic Action of Apes. London: Allison & Busby, 1980; revised edition.

Wilson, D. R. (1993).
Evolutionary epidemiology: Darwinian theory in the service of medicine and psychiatry. *Acta Biotheoretica,* Vol. 41, pp. 205-218.

Wright, R. (1994).
The Moral Animal: The New Science of Evolutionary Psychology. New York: Vintage Books.

University of California, Berkeley
Department of Music

The 1999 Ernest Bloch Lectures

Lecture 3. Methodology:

The New Empiricism: Systematic Musicology in a Postmodern Age

David Huron

KEYWORDS: methodology, empiricism, postmodernism, musicology

ABSTRACT: A survey of intellectual currents in the philosophy of knowledge and research methodology is given. This survey provides the backdrop for taking stock of the methodological differences that have arisen between disciplines, such as the methods commonly used in science, history or literary theory. Postmodernism and scientific empiricism are described and portrayed as two sides of the same coin we call skepticism. It is proposed that the choice of methodological approach for any given research program is guided by moral and esthetic considerations. Careful assessment of these risks may suggest choosing an unorthodox method, such as quantitative methods in history, or deconstruction in science. It is argued that methodological tools (such as Ockham's razor) should not be mistaken for philosophical world-views. The article advocates a broadening of methodological education in both arts and sciences disciplines. In particular, it advocates and defends the use of quantitative empirical methodology in various areas of music scholarship.

Introduction

[1] Scholarly disciplines distinguish themselves from one another, principally by their subject matter. Musicology differs from chemistry, and chemistry differs from political science because each of these disciplines investigates different phenomena. Apart from the subject of study, scholarly disciplines also frequently differ in how they approach research. The methods of the historian, the scientist, and the literary scholar often differ dramatically. Moreover, even within scholarly disciplines, significant methodological differences are common.

[2] Over the past two decades, music scholarship has been influenced by at least two notable methodological movements. One of these is the so-called "new musicology." The new musicology is loosely guided by a recognition of the limits of human understanding, an awareness of the social milieu in which scholarship is pursued, and a realization of the political arena in which the fruits of scholarship are used and abused. The influence of the new musicology is evident primarily in recent historical musicology and ethnomusicology, but it has proved broadly influential in all areas of music scholarship, including music education.

[3] Simultaneously, the past two decades have witnessed a rise in scientifically inspired music research. This increase in empirical scholarship is apparent in the founding of several journals, including *Psychomusicology* (founded 1981), *Empirical Studies in the Arts* (1982), *Music Perception* (1983), *Musicae Scientiae* (1997), and *Systematic Musicology* (1998). This new empirical enthusiasm is especially evident in the psychology of music and in the resurrection of systematic musicology. But empiricism is also influential in certain areas of music education and in performance research. Music researchers engaged in empirical work appear to be motivated by an interest in certain forms of rigor, and a belief in the possibility of establishing positive, useful musical knowledge.

[4] The contrast between the new musicology and the new empiricism could hardly be more stark. While the new musicology is not merely a branch of Postmodernism, the influence of Postmodern thinking is clearly evident. Similarly, while recent music empiricism is not merely the offspring of Positivism, the family resemblance is unmistakable. Yet the preeminent intellectual quarrel of our time is precisely that between Positivism and Postmodernism -- two scholarly approaches that are widely regarded as mortal enemies. {1} How have these diametrically opposed methodologies arisen, and what is a thoughtful scholar to learn from the contrast? How indeed, ought one to conduct music research?

[5] By *methodology*, I mean any formal or semi-formal approach to acquiring insight or knowledge. A methodology may consist of a set of fixed rules or injunctions, or it may consist of casual guidelines, suggestions or heuristics. From time to time, a particular methodology emerges that is shared in common by several disciplines. One example is the so-called Neyman-Pearson paradigm for inductive empirical research commonly used in the physical sciences (Neyman and Pearson, 1928, 1967). But not all disciplines adopt the same methodologies, nor should they.

[6] Different research goals, different fears, different opportunities, and different dispositions can influence the adoption and development of research methods. For any given scholarly pursuit, some research methods will prove to be better suited than others. Part of the scholar's responsibility then, is to identify and refine methods that are appropriate to her or his field of study. This responsibility includes recognizing when a popular research method ceases to be appropriate, and adapting one's research to take advantage of new insights concerning the conduct of research as these insights become known.

Two Cultures

[7] Historically, the most pronounced methodological differences can be observed in the broad contrast between the sciences and the humanities. (For convenience, in this article I will use the term "humanities" to refer to both the humanities and the arts.) In humanities scholarship, research methods include historiographic, semiotic, deconstructive, feminist, hermeneutical, and many other methods. In the sciences, the principal scholarly approaches include modeling and simulation, analysis-by-synthesis, correlational and experimental approaches.

[8] Many scholars presume that methodological differences reflect basic philosophical disagreements concerning the nature of scholarly research. I think this view masks the more fundamental causes of methodological divergence. As I will argue in this article, in most cases, the main methodological differences between disciplines can be traced to the materials and circumstances of the particular field of study. That is, differences in research methods typically reflect concrete differences between fields (or sub-fields) rather than reflecting some underlying difference in philosophical outlook. This is the reason, I will contend, why Muslims and Christians, atheists and anarchists, liberals and libertarians, have little difficulty working with each other in most disciplines. Although deep personal beliefs may motivate an individual to work on particular problems, one's core philosophical beliefs often have little to do with one's scholarly approach.

Philosophy of Knowledge and Research Methodology

[9] In addressing issues pertaining to scholarly methodology, there is merit in dividing the discussion into two related topics. One topic relates to broad epistemological issues, while the second topic relates to the concrete

issues of how one goes about doing practical scholarship. In short, we might usefully distinguish philosophy of knowledge (on the one hand) from research methodology (on the other). One rightly expects that the positions we hold regarding the philosophy of knowledge would inform and shape the concrete procedures we use in our day-to-day research methods. However, the information flows in both directions. Practical research experiences also provide important lessons that shape our philosophies of knowledge.

[10] In the training of new scholars, it appears that academic disciplines often differ in the relative weight given to philosophy of knowledge compared with research methodology. My experience with psychologists, for example, is that they typically receive an excellent training in the practical nuts and bolts of research methodology. In conducting research, there are innumerable pitfalls to be avoided, such as *confirmation bias*, *demand characteristics*, and *multiple tests*. These are the sorts of things experimental psychologists learn to recognize, and devise strategies to avoid or minimize. However, most psychologists I have encountered have received comparatively less training in the philosophy of knowledge. Most have only heard of Hume and Popper, van Quine and Lakatos, Gellner, Laudan, and others. The contrast with the training of literary scholars is striking. There is hardly an English scholar, trained in recent decades, who has not read a number of books pertaining to the philosophy of knowledge. The list of authors differs, however -- emphasizing the anti-foundationalist writers: Kuhn and Feyerabend, Derrida and Foucault, Lacan, Leotard, and others. {2} On the other hand, most English scholars receive relatively little training in research methodology, and this is often evident in the confusion experienced by young scholars when they embark on their own research: they often don't know how to begin or what to do.

[11] The philosophical and methodological differences between the sciences and the humanities can be the cause of considerable discomfort for those of us working in the gap between them. As a cognitive musicologist, I must constantly ask whether I should study the musical mind as a humanities scholar, or as a scientist? Having given some thought to methodological questions, my purpose in this article is to share some observations about these convoluted yet essential issues.

Overview

[12] My goal in this article is to take stock of the methodological differences that arise between disciplines and to attempt to understand their origins and circumstantial merits. As I've already noted, I think the concrete circumstances of research are especially formative. However, before I argue this case, it behooves me to address the noisy (and certainly interesting) debates in the philosophy of knowledge. In particular, it is appropriate to address the often acrimonious debate between empiricism and postmodernism.

[13] Of course not all sciences are empirical and not all humanities scholarship is postmodern. The field of mathematics (which is popularly often considered "scientific") relies almost exclusively on deductive methods rather than empirical methods. Similarly, although postmodernism has been a dominant paradigm in many humanities disciplines over the past two decades, there exist other methodological traditions in humanities scholarship. The reason why I propose to focus on the empirical and postmodernist traditions is that they are seemingly the most irreconcilable. I believe we have the most to learn by examining this debate.

[14] This paper is divided into two parts. In Part I, I outline some of the intellectual history that forms the background for contemporary empiricism and postmodernism. Part II focuses more specifically on methodology. In particular, I identify what I think are the principal causes that lead to the adoption of different methodologies in different fields and sub-fields. Part II also provides historical examples where disciplines have dramatically changed their methodological preferences in response to new circumstances. My claim is that the resources available for music scholarship are rapidly evolving, and that musicology has much to gain by adapting empirical methods to many musical problems. I conclude by outlining some of the basic ideas underlying what might be called the "new empiricism."

PART ONE: Philosophy of Knowledge

Empiricism and Science

[15] The dictionary definition of "empirical" is surprisingly innocuous for those of us arts students who were taught to use it as a term of derision. Empirical knowledge simply means knowledge gained through observation. Science is only one example of an empirical approach to knowledge. In fact, many of the things traditional historical musicologists do are empirical: deciphering manuscripts, studying scores, and listening to performances.

[16] The philosophical complexity begins when one asks how it is that we learn from observation. The classic response is that we learn through a process dubbed *induction.* Induction entails making a set of specific observations, and then forming a general principal from these observations. For example, having stubbed my toe on many occasions over the course of my life, I have formed a general conviction that rapid movement of my toe into heavy objects is likely to evoke pain. We might say that I have learned from experience (although my continued toe-stubbings make me question how well I've learned this lesson).

[17] The 18th-century Scottish philosopher, David Hume, recognized that there are serious difficulties with the concept of induction. Hume noted that no amount of observation could ever resolve the truth of some general statement. For example, no matter how many white swans one observes, an observer would never be justified in concluding that *all swans are white.* Using postmodernist language, we would say that one cannot legitimately raise local observations to the status of global truths.

[18] Several serious attempts have been made by philosophers to resolve the problem of induction. Three of these attempts have been influential in scientific circles: *falsificationism, conventionalism* and *instrumentalism*. However these attempts suffer from serious problems of their own. In all three philosophies, the validity of empirical knowledge is preserved by forfeiting any strong claim to absolute truth.

[19] One of the most influential epistemologies in twentieth-century empiricism was the philosophy of *conventionalism*. The classic statement is found in Pierre Duhem's *The Aim and Structure of Physical Theory* originally published in 1905, but reprinted innumerable times throughout the past century. In his book, Duhem notes that science never provides theories or explanations of some ultimate reality. Theoretical entities and mathematical laws are merely conventions that summarize certain types of relationships. It can never be determined whether scientific theories are "true" in the sense of explaining or capturing some underlying reality. Scientific theories are merely conventions that help scientists organize the observable patterns of the world.

[20] A variation of conventionalism, known as *instrumentalism* similarly posits that empiricism does not provide ultimate explanations: the engineer has no deep understanding of why a bridge does not fall down. Rather, the engineer relies on theories as tools that are reasonably predictive of practical outcomes. For the instrumentalist, theories are judged, not by their "truthfulness," but by their predictive utility.

[21] The most well-known attempt to resolve the problem of induction was formulated by Karl Popper in 1934. Popper accepted that no amount of observation could ever verify that a particular proposition is true. That is, an observer cannot prove that *all swans are white*. However, Popper argued that one could be certain of falsity. For example, observing a single black swan would allow one to conclude that the claim -- *all swans are white* -- is false. Accordingly, Popper endeavored to explain the growth of knowledge as arising by trimming the tree of possible hypotheses using the pruning shears of falsification. Truth is what remains after the falsehoods have been trimmed away.

[22] Popper's approach was criticized by van Quine, Lakatos, Agassi, Feyerabend and others. One problem is that it is not exactly clear what is falsified by a falsifying observation. It may be that the observation itself is incorrect, or the manner by which the phenomenon of interest is defined, or the overall theoretical framework within which a specific hypothesis is posited. (For example, the observer of a purported black swan might have

been drunk, or the swan might have been painted, or the animal might be claimed to be a different species.) A related problem is fairly technical, and so difficult to describe succinctly. In order to avoid prematurely jettisoning a theory, Popper abandoned the notion of a *falsifying observation* and replaced it with the concept of a *falsifying phenomenon*. Yet to establish a falsifying phenomenon, researchers must engage in an activity of verification -- an activity which Popper himself argued was impossible. In Popper's methodology, the nasty problem of inductive truth returns through the rear door.

[23] Despite such difficulties, Popper's falsificationism has remained highly influential in the day-to-day practice of empirical research. In the professional journals of science, editors regular remove claims that *such-and-such is true*, or that *such-and-such a theory is verified*, or even that *the data "support" such-and-such a hypothesis*. On the contrary, the boiler-plate language for scientific claims is: *the null hypothesis was rejected* or *the data are consistent with such-and-such a hypothesis.* Of course this circumspect language is abandoned in secondary and popular scientific writings, as well as in the informal conversations of scientists. This gap between official skepticism and colloquial certainty is a proper subject of study for sociologists of science.

[24] Another, less influential scientific epistemology in the twentieth century was *positivism.* Positivism never provided a proposal for resolving the problem of induction. Nevertheless, it is worth brief mention here for two reasons. First logical positivism drew attention to the issue of language and meaning in scientific discourse, and secondly, "positivism" has been the preeminent target of postmodernist critiques.

[25] Positivism began as a social philosophy in France, initiated by Saint-Simon and Comte, and spread to influence the sciences in the early twentieth century. The tenants of positivism were articulated by the so-called Vienna Circle (including Schlick and Carnap) and culminated in the classic statement of 1936 by A.J. Ayer. In science, logical positivism held sway from roughly 1930 to 1965. However, this influence was almost exclusively restricted to American psychology; only a small minority of empiricists ever considered themselves positivists.

[26] For most of the twentieth century, the preeminent philosophical position of practicing scientists (at least those scientists who have cared to comment on such matters) has been conventionalism or instrumentalism. Popper's emphasis on falsifying hypotheses (which is consistent with both conventionalism and instrumentalism) has proved highly influential in the day-to-day practice of science, largely because of the Pearson/Neyman/Popper statistically-based method of inductive falsification. (Many epistemologists consider Popper's most important and influential writings to be his appendices on probability and statistics.)

[27] This is by no means a complete story of the philosophy of science in the twentieth century, but before we continue our story, it is appropriate to turn our attention to postmodernism.

Postmodernism

[28] Postmodernism is many things, and any attempt to summarize it is in danger of oversimplification. (Indeed, one of the principal tenants of postmodernism is that one should not attempt to represent the world-views of others.) In the same way that philosophers of science disagree with one another, those who call themselves postmodernists also are not of one mind. Nevertheless, there are a number of common themes that tend to recur in postmodernist writings. Postmodernism is a philosophical movement that focuses on how meanings get constructed, and how power is commandeered and exercised through language, representation and discourse. {3}.

[29] Postmodernism is interested in scholarship, because scholarly endeavors are among the preeminent meaning-conferring activities in our society. Postmodernism is especially interested in science, principally because, at least in Western societies, science holds a power of persuasion second to no other institution. It is a power, of which the most powerful politicians can only express envy.

[30] Postmodernism begins from a position surprisingly similar to Popper's anti-verification stance and Duhem's conventionalism. Where Duhem and Popper thought that the truth is unknowable, postmodernism assumes that

there is no absolute truth to be known. More precisely, "truth" ought to be understood as a social construction that relates to a local or partial perspective on the world. Our mistake is to assume that as observers, we can climb out of the box which is our world. There is no such objective perspective.

[31] There are, rather, a vast number of interpretations about the world. In this, the world is akin to a series of texts. As illustrated in the writings of Jacques Derrida, any text can be deconstructed to reveal multiple interpretations, no one of which can be construed as complete, definitive, or privileged. From this, postmodernists conclude that there is no objective truth, and similarly that there is no rational basis for moral, esthetic or epistemological judgment.

[32] If there is no absolute basis for these judgments, how do people in the world go about making the decisions they do? The most successful achievements of postmodernism have been in drawing attention to the power relations that exist in any situation where an individual makes some claim. As Nancy Hartsock has suggested, "the will to power [is] inherent in the effort to create theory" (1990; p.164). Like the politician or the business person, scholars are consciously or unconsciously motivated by the desire to commandeer resources and establish influence. Unlike the politician or the business person, we scholars purport to have no hidden agenda -- a self-deception that makes us the most dangerous of all story-tellers.

[33] It is the most powerful members of society who are able to establish and project their own stories as so-called "master narratives." These narratives relate not only to claims of truth, but also to moral and artistic claims. The "canons" of art and knowledge are those works exalted by, and serving, the social elites. Insofar as works of art give legitimacy to those who produce them, "A work of art is an act of power." (Rahn, 1993)

[34] This admittedly pessimistic view of the world could well lead one to despair. Since there is no legitimate power, how does the conscientious person act so as to construct a better world? Postmodernism offers various strategies that might be regarded as serving the goal of *exposé*. That is, the postmodernist helps the cause through a sort of investigative journalism that exposes how behaviors are self-serving. At its best, postmodernism is a democratizing ladle that stirs up the political soup and resists the entrenchment of a single power. By creating a sort of chaos of meaning, it calls existing canons into question, subverts master narratives, and so gives flower to what has been called "the politics of difference".

Feyerabend and the Galileo-Scholastics Debate

[35] In the world of the sciences, a concrete demonstration of such power relations is examined in the work of Paul Feyerabend. In his book, *Against Method.* Feyerabend used scientific method itself to show the failures of scientific discourse, and the role of power in presumed rational debate.

[36] It is worth discussing Feyerabend's work at some length because his work has led to widespread misconceptions, many of which were promoted by Feyerabend himself.

[37] Contemporary scientific method embraces certain standards for evidence in scientific debates. For example, when two competing theories (*X* and *Y*) exist, scientists attempt to construct a "critical experiment" where the two theories are pitted against each other. If the results turn out one way, theory *X* is rejected; if the results turn out another way, theory *Y* is rejected. In addition, contemporary scientific method frowns upon so-called *ad hoc* hypotheses. Suppose that the results of a critical experiment go against my pet theory. I might try to save my theory by proposing that the experiment was flawed in various ways. I might say that the reason the experiment failed to be consistent with my theory is that the planet Mercury was in retrograde on the day that the experiment was carried out, or that my theory is true except on the third Wednesday of each month. Of course *ad hoc* hypotheses need not be so fanciful. More credible *ad hoc* hypotheses might claim that the observer was poorly trained, the equipment not properly calibrated, or the control group improperly constructed, etc. Although an *ad hoc* hypothesis might be true, such appeals are considered very bad form in scientific circles whenever the motivation for such claims is patently to "explain away" a theoretical failure.

[38] Feyerabend uses the case study of the famous debate between Galileo and the Scholastics. In the popular understanding of this history, Galileo argued that the sun was positioned in the center of the solar system and the Scholastics, motivated by religious dogma, maintained that the earth was in the center of the universe.

[39] Historically, this popular view is not quite right -- as Feyerabend points out. The Scholastics argued that motion is relative, and that there is, in principle, no way that one could determine whether the earth was rotating about the sun or the sun was rotating about the earth. Since observation alone cannot resolve this question, the Scholastics argued that the Bible implies that the earth would be expected to hold a central position.

[40] However, Galileo and the Scholastics agreed on a possible critical experiment. Suppose that your head represents the earth. If you rotate your head in a fixed position, the angles between various objects in the room will remain fixed. However, if you walk in a circle around the room, the visual angles between various objects will change. As you approach two objects, the angle separating them will increase. Conversely, as you move away from two objects, the angle separating them will decrease.

[41] According to this logic, if the earth is in motion, then one ought to be able to see slight angular shifts between the stars over the course of the year. Using his new-fangled invention, the telescope, Galileo did indeed make careful measurements of the angular relationships between the stars over the course of a year. He found, however, that there was no change whatsoever. In effect, Galileo carried out a critical experiment -- one whose results were not consistent with the idea that the earth is in motion. How did Galileo respond to this result? Galileo suggested that the reason why no parallax shifts could be observed was because the stars are extremely far away.

[42] Feyerabend pointed out that this is an *ad hoc* hypothesis. A critical experiment was carried out to determine whether the earth or the sun was in motion, and Galileo's theory lost. Moreover, Galileo had the audacity to defend his theory by offering an *ad hoc* hypothesis. By modern scientific standards, one would have to conclude that the Scholastics' theory was superior, and that, as a scientist, Galileo himself should have recognized that the evidence was more consistent with the earth-centered theory.

[43] Of course, from our modern perspective, Galileo was right to persevere with his sun-centered theory of the solar system. As it turns out, his *ad hoc* hypothesis regarding the extreme distance to the stars is considered by astronomers to be correct.

[44] From this history, Feyerabend draws the following conclusions. First, the progress of science may depend on bad argument and ignoring data. Second, Galileo should be recognized, not as a great scientist, but as a successful propagandist. Third, had Galileo followed modern standards of scientific method the result would have been scientifically wrong. Fourth, the injunction against *ad hoc* hypotheses in science can produce scientifically incorrect results. Fifth, the use of critical experiments in science can produce scientifically incorrect results. Sixth, no methodological rule will ensure a correct result. Seventh, there is no scientific method. And eighth, in matters of methodology, concludes Feyerabend, *anything goes.* Like Popper and Lakatos, Feyerabend argued that there is no set of rules that guarantees the progress of knowledge.

[45] In assessing Feyerabend's work, we need to look at both his successes and failures. Let's begin with some problems. Recall that the problem of induction is the problem of how general conclusions can be drawn from a finite set of observations. Consider, the fourth and fifth of Feyerabend's conclusions. He notes that two rules in scientific methodology (namely, the rule forbidding *ad hoc* hypotheses, and the instruction to devise critical experiments) failed to produce a valid result in Galileo's case. From these two historical observations, Feyerabend formulates the general conclusion: no methodological rule will ensure a correct result. By now you should recognize that this is an inductive argument, and as Hume pointed out, we can't ever be sure that generalizing from specific observations produces a valid generalization.

[46] Showing that *some* methodological rules don't work in a single case, doesn't allow us to claim that all methodological rules are wrong. Even if one were to show that *all* known methodological rules were inadequate, one can't logically conclude than there are *no* true methodological rules.

[47] A further problem with Feyerabend's argument is that he exaggerates Galileo's importance in the promotion of the sun-centered theory. The beliefs and arguments of a single person are typically limited. Knowledge is socially distributed, and ideas catch on, only when the wider population is prepared to be convinced. In fact, the heliocentric theory of the solar system was not immediately adopted by scientists because of Galileo's arguments. The heliocentric theory didn't gain many converts until after Kepler showed that the planets move in elliptical orbits. Kepler's laws made the sun-centered theory a much simpler system for describing planetary motions. In short, Galileo's fame and importance as a scientific champion is primarily retrospective and ahistorical.

[48] Feyerabend's historical and analytic work is insufficient to support his general conclusion: namely that in methodology, the only correct rule is "anything goes." Moreover, Feyerabend's own dictum is not born out by observation. Anyone observing any meeting of any academic group will understand that, in their debates, it is not true that `anything goes.' All disciplines have more or less loose standards of evidence, of sound argument, and so on. Although a handful of scholars might wish that debates could be settled through physical combat, for the majority of scholars such "methods" are no longer admissible. There may be no methodological recipe that guarantees the advance of knowledge, but similarly, it is not the case that anything goes.

[49] On the positive side, Feyerabend has drawn attention to the social and political environment in which science takes place. Feyerabend stated that his main reason for writing *Against Method* was "humanitarian, not intellectual". Feyerabend wanted to provide rhetorical support for the marginalized and dispossessed (p.4). In drawing attention to the sociology of science, Feyerabend and his followers have met strong resistance from scientists themselves. Until recently, most scientists rejected the notion that science is shaped by a socio-political context. The failings of science notwithstanding, this does not mean that scholars working in the sociology of science have been doing a good job.

Kuhn and Paradigmatic Research

[50] The most influential study of science is probably Thomas Kuhn's *The Structure of Scientific Revolutions*. As a historian of science, Kuhn set out to describe how new ideas gain acceptance in a scientific community.

[51] From his studies in the history of science Kuhn distinguished two types of science: *normal* science and *revolutionary* science. The majority of scientific research can be described as normal science. Normal science is a sort of puzzle-solving activity, where the prevailing scientific theory is applied in various tasks, and small anomalies in the prevailing theory are investigated. Many anomalies are resolved by practicing such "normal" science. However, over time, certain anomalies fail to be resolved and a minority of scientists begin to believe that the prevailing scientific theory (or "paradigm") is fundamentally flawed.

[52] *Revolutionary science* breaks with the established paradigm. It posits an alternative interpretation that meets with stiff resistance. Although the new theory might explain anomalies in the prevailing theory, inevitably, there are may things that are not (yet) accounted for by the new theory. Opponents of the new paradigm contrast these failures with the known successes of the existing paradigm. (In part, the problems with the new paradigm can be attributed to the fact that the new theory has not yet benefitted from years of normal science that resolve apparent problems that can be explained using the old paradigm.)

[53] An important claim made by Kuhn is that debates between supporters of the old and new paradigms are not rational debates. Changing paradigms is akin to a religious conversion: one either sees the world according to the old paradigm or according to the new paradigm. Supporters of the competing paradigms are incapable of engaging each other in reasoned discussion. Scientists from competing paradigms "talk past each other." Technical terms, such as "electron" begin to have different meanings for scientists supporting different paradigms.

[54] Kuhn argued that there is no neutral or objective position from which one can judge the relative merits of the two different paradigms. Consequently, Kuhn characterized the paradigms as *incommensurable* -- not measurable using a single yard-stick. Paradigm shifts occur, not because supporters of the old paradigm become

convinced by the new paradigm. Instead, argues Kuhn, new paradigms replace old paradigms because old scientists die, and new paradigm supporters are able to place their colleagues and students in important positions of power (professorships, journal editors, granting agencies, etc.) Once advocates of the new paradigm have seized power, the textbooks in the discipline are re-written so that the revolutionary change is re-cast as a natural and inevitable step in the continuing smooth progress of the discipline.

[55] While Kuhn's work had an enormous impact in the social sciences, it had comparatively little impact in the sciences themselves. *The Structure of Scientific Revolutions* portrayed science as akin to fashion: changes do not arise from some sort of rational debate. Change is simply determined by who holds power. Although Thomas Kuhn denied that he was arguing that science does not progress, his study of the history of science strongly implies that "scientific progress" is an illusion perpetrated by scientists who re-construct history to place themselves (and their paradigms) at the pinnacle of a long lineage of achievement.

[56] Many social sciences and humanities scholars applauded Kuhn because his portrayal removed science from the epistemological high ground. The presumed authority of science is unwarranted. Like different cultures around the world, there is no valid yard-stick by which one can claim that one scientific culture is better than another.

[57] Kuhn's writings also appealed to those scientists (and other scholars) whose views place them outside the mainstream. For those scientists whose unorthodox views are routinely ignored by their colleagues, Kuhn's message is highly reassuring. The reason why other people don't understand us and don't care about what we say, is that they are enmeshed in the old paradigm: no amount of reasoned debate can be expected to convince the existing powers. In short, Kuhn's characterization of science provides a measure of comfort to the marginalized and dispossessed.

[58] Shortly after the publication of Kuhn's book, a young Bengali philosopher named Jagdish Hattiangadi wrote a detailed critique of the work. Although Kuhn regarded himself as a historian of science with great sympathies for science, Hattiangadi noted that Kuhn's work removed any possibility that science could be viewed as a rational enterprise. Although Kuhn never said as much, his theory had significant repercussions: for example, a chemist who believes that modern chemistry is *better* than ancient chemistry must simply be deluded. Hattiangadi noted that, either there is no progress whatsoever in science, or Kuhn's portrayal of science is wrong. Hattiangadi concluded that Kuhn's work failed to account for the widespread belief that scientific progress is a fact. Moreover, as early as 1963, Hattiangadi predicted that Kuhn's book would become wildly successful among social and humanities scholars -- a prediction that proved correct.

Postmodernism: An Assessment

[59] With this background in place, let's return to our discussion of postmodernism. In general, postmodernism takes issue with the Enlightenment project of deriving absolute or universal truths from particular knowledge. That is, postmodernism posits a radical opposition to induction. We cannot generalize from the particular; the global does not follow from the local.

[60] At first glance, it would appear that postmodernism would be as critical of Feyerabend and Kuhn as of the positivists. For the arguments of Feyerabend and Kuhn also rest on the assumption that we can learn general lessons from specific historical examples. However, postmodernism is less concerned with such convoluted issues than it is with the general goal of causing intellectual havoc for those who want to make strong knowledge claims. Accordingly, the works of Feyerabend and Kuhn are regarded as allies in the task of unraveling science's presumed authority.

[61] Of course postmodernism also has its critics. Much of the recent unhappiness with postmodernism is that it appears to deny the possibility for meaningful human change. For example, many feminist thinkers have dismissed a postmodernist approach because it removes the high moral ground. In lobbying for political change, most feminists have been motivated by a sense of injustice. However, if there are no absolute precepts of justice, then the message postmodernism gives to feminists is that they are simply engaged in Machiavellian maneuvers

to wrest power. In the words of Joseph Natoli, "postmodernist politics here has nothing to do with substance but only with the tactics." (1997, p. 101) On the one hand, postmodernism encourages feminists to wrest power away from the male establishment; but at the same time, postmodernism tells feminists not to believe that their actions are at all justified. Understandably, many feminists are uncomfortable with this contradiction.

[62] The nub of the issue, I think, is evident in the following two propositions associated with postmodernism:

> (1) There is no privileged interpretation.
> (2) All interpretations are equally valid.

As the postmodernist writer Catherine Belsey has noted, postmodernism has been badly received by the public primarily because postmodernists have failed to distinguish between sense and nonsense. This is the logical outcome for those who believe that (2) is simply a restatement of (1).

[63] If we accept the proposition that there is no privileged interpretation, it does not necessarily follow that all interpretations are equally valid. For those who accept (1) but not (2), it follows that some interpretations must be "better" than others -- hence raising the question of what is meant by "better."

[64] Postmodernism has served an important role by encouraging scholars to think carefully, laterally, and self-reflectively. Unfortunately, postmodernism encourages slovenly research and a disinterest in pursuing rigor. Postmodernism draws welcome attention to the social and political context of knowledge and knowledge claims. But postmodernism goes too far when it concludes that reality is socially *constructed* rather than socially *mediated*. Postmodernism serves an important role when it encourages us to think about power relations, and in particular how certain groups are politically disenfranchised because they have little control over how meanings get established. But at the same time, postmodernism subverts all values, and transforms justice into mere tactical maneuvers to gain power. In reducing all relationships to power, postmodernism leaves no room for other human motivations. Scholarship may have political dimensions, but that doesn't mean that all scholars are plotting power-mongers. Postmodernism is important insofar as it draws attention to the symbolic and cultural milieu of human existence. But, while we should recognize that human beings are cultural entities, we must also recognize that humans are also biological entities with *a priori* instinctive and dispositional knowledge about the world that originates in an inductive process of evolutionary adaptation (Plotkin, 1994). Foucault regrettably denied any status for humans as biological entities whose mental hardware exists for the very purpose of gaining knowledge about the world.

[65] When pushed on the issue of relativism, postmodernists will temporarily disown their philosophy and accept the need for some notion of logic and rigor. Belsey, for example, claims that as postmodernists, "we should not abandon the notion of rigor; the project of substantiating our readings" (Belsey, 1993, p. 561) Similarly, Natoli recognizes that "logic" (1997, p.162) and "precision" (p.120) make for compelling narratives. However, postmodernists are oddly uninterested in how these approaches gain their rhetorical power. What is "logic"? What is "rigor"? What is it about rationality that makes some narratives so mentally seductive or compelling? It is exactly this task that has preoccupied philosophers of knowledge over the past 2,500 years and was the focus of Enlightenment efforts in epistemology. The Enlightenment project of attempting to characterize the value of various knowledge claims is not subverted by postmodernism. On the contrary, postmodernism simply raises anew the question of what it means to do good scholarship.

PART TWO: Philosophy of Methodology

[66] How then, should scholars conduct research? What does the philosophy of knowledge tell us about the practicalities of scholarship? As we have seen, the philosophy of knowledge suggests that we abandon the view

that methodology is an infallible recipe or algorithm for establishing the truth. The epistemological role of methodology is much more modest. At the same time, what the new empiricism shares in common with postmodernism is the conviction that scholarship occurs in a moral realm, and so methodology ought be guided by moral considerations.

Methodological Differences

[67] As noted in the introduction, one of the principal goals of this paper is to better account for why methodologies differ for different disciplines. In pursuing this goal I will outline a taxonomy of research methodologies based on four distinctions. In brief, these are:

- *False-positive skepticism versus false-negative skepticism.* False-positive skepticism holds that theories or hypotheses ought to be rejected given the slightest contradicting evidence. False-negative skepticism holds that theories or hypotheses ought to be conserved unless there is overwhelming contradicting evidence.
- *High risk versus low risk theories.* Theories, hypotheses, interpretations and intuitions carry moral and esthetic repercussions. In testing some knowledge claim, the burden of evidence can shift depending on the consequences of the theory. Many theories carry negligible risks, however.
- *Retrospective versus prospective data.* Some areas of research (such as manuscript studies) have only pre-existing evidence or data. Other areas of research (such as behavioral studies) have opportunities to collect newly generated evidence. Prospective data allows researchers to more rigorously test knowledge claims by attempting to forecast properties of yet-to-be-collected data.
- *Data-rich versus data-poor fields.* Fields of study can also be characterized according to the volume of pertinent evidence. When the evidence is minimal, researchers in data-rich fields have the luxury of suspending judgment until more evidence is assembled. By contrast, researchers in data poor fields often must interpret a set of data that is both very small and final -- with no hope of additional forthcoming evidence.

[68] Below, I will describe more fully these four distinctions. My claim is that fields of study can be usefully characterized by these taxonomic categories. Each of these four distinctions has repercussions for formulating field-appropriate methodologies. I will suggest that these taxonomic distinctions not only help us to better understand why methodologies diverge for various fields, but also help us to better recognize when an existing methodology is inappropriate for some area of study.

[69] Additionally, I will note that fields of research sometimes experience major changes in their basic working conditions -- changes that precipitate shifts in methodology. A formerly uncontentious field of research (such as education) may abruptly find that its latest theories {4} carry high moral risk. A previously data-poor field (such as theology) may become inundated by new sources of information. And a formerly retrospective discipline (such as history) may unexpectedly find a class of events for which it can offer testable predictions. Later in this article I will briefly discuss two case examples of such shifts in resources and methods. My first example is the transformation of sub-atomic physics so that its methods increasingly resemble those in philosophy and literary theory. My second example will be the increasing influence of empirical methods in music scholarship.

Two Forms of Skepticism

[70] From at least the time of the ancient Greeks, the essence of scholarship has been closely associated with skepticism. Most scholars evince a sort of love/hate relationship with skepticism. On the one hand, we have all experienced annoyance at the credulity of those who accept uncritically what we feel ought to evoke wariness. On the other hand, we have all experienced exasperation when someone offers belligerent resistance to the seemingly obvious. What one person regards as prudent reserve, another considers bloody-mindedness.

[71] Science is often portrayed as an institutionalized form of skepticism. Unfortunately, this portrayal can leave the false impression that the arts and humanities are *not* motivated by skepticism -- that the humanities are somehow credulous, doctrinaire, or gullible. Contrary to the views of some, most humanities disciplines also

cultivate institutionalized forms of skepticism; however, the type of skepticism embraced is often diametrically opposed to what is common in the sciences.

[72] These differences are illustrated in Table 1. The table identifies four epistemological states related to any knowledge claim (including the claim that something is unknowable). Whenever a claim, assertion, or mere insinuation is made, two types of errors are possible. A *false positive error* occurs when we claim something to be true or useful or knowable when it is, in fact, false, useless or unknowable. A *false negative error* occurs when we claim something to be false/useless/unknowable when it is, in fact, true/useful/knowable. Methodologists refer to these errors as Type I and Type II respectively.

Table 1

	Thought to be True, Useful or Knowable	**Thought to be False, Useless or Unknowable**
Actually True, Useful or Knowable	Correct Inference	False Negative Error (Type II Error)
Actually False, Useless or Unknowable	False Positive Error (Type I Error)	Correct Inference

[73] The *false-positive skeptic* tends to make statements such as the following:

> "You don't know that for sure."
> "I really doubt that that's useful."
> "There's no way you could ever know that."

By contrast, *false-negative skepticism* is evident in statements such as the following:

> "It might well be true."
> "It could yet prove to be useful."
> "We might know more than we think."

In short, the two forms of skepticism might be summarized by the following contrasting assertions:

> *False-Positive Skeptic:* "There is insufficient evidence to support that."
> *False-Negative Skeptic:* "There is insufficient evidence to reject that."

[74] Speaking of false-negative and false-positive skepticism can be a bit confusing. For the remainder of this article, I'll occasionally refer to false-positive skepticism as *theory-discarding skepticism* since these skeptics look for reasons to discard claims, theories or interpretations. By contrast, I'll occasionally refer to false-negative skepticism as *theory-conserving skepticism* since these skeptics are wary of evidence purporting to disprove a theory or dismiss some claim, view, interpretation or intuition.

[75] In the case of the physical and social sciences, most researchers are theory-discarding skeptics. They endeavor to minimize or reduce the likelihood of making false-positive errors. That is, traditional scientists are loath to make the mistake of claiming something to be true that is, in reality, false. Hundreds of thousands of scientific publications begin from the premise of theory-discarding skepticism. {5} This practice has arisen in response to researchers' observations that we are frequently wrong in our intuitions and all too eager to embrace suspect evidence in support of our pet theories.

[76] In the past two decades or so, medical researchers have raised serious challenges to this orthodox scientific position. The U.S. Food and Drug Administration formerly approved only those drugs that had been proved to be effective (i.e., "useful") according to criteria minimizing false-positive errors. (That is, drugs that *might* be useful were rejected.) The AIDS lobby drew attention to the illogic of denying seemingly promising drugs that had not yet been shown to be useless. For the patient facing imminent death, it is the enlightened physician who will recommend that her patient seek out the most promising of recent "quacks." {6}. In other words, the medical community has drawn attention to the possible detrimental effects of committing false-negative errors. Theory-discarding skeptics are prone to the error of claiming something to be useless that is, in fact, useful.

[77] This shift in attitude has moved contemporary medical research more closely towards dispositions more commonly associated with traditional arts/humanities scholars. Broadly speaking, traditional humanities scholars (including scholars in the arts) have tended to be more fearful of committing false-negative errors. For many arts and humanities scholars, a common fear is dismissing prematurely an interpretation or theory that might have merit -- however tentative, tenuous or incomplete the supporting evidence. Arts scholars (in particular) have placed a premium on what is regarded as sensitive observation and intuition: no detail is too small or too insignificant when describing or discussing a work of art.

[78] Another way that traditional humanities scholars exhibit theory-conserving tendencies is evident in attitudes toward the notion of *coincidence*. For traditional scientists, the principal methodological goal is to demonstrate that the recorded observations are unlikely to have arisen by chance. In the common Neyman-Pearson research paradigm, this is accomplished by disconfirming the *null hypothesis*. That is, the researcher makes a statistical calculation showing that the observed data {7}. are inconsistent with the hypothesis that the data would be expected to arise by chance. For many traditional humanities scholars, however, dismissing an observation as a "mere coincidence" is problematic. If the goal is to minimize false negative claims, then a single "coincidental" observation should not be dismissed lightly. For many arts and humanities scholars, apparent coincidences are more commonly viewed as "smoking guns."

[79] In summary, both traditional scientists and traditional humanities scholars are motivated by skepticism, but they often appear to be motivated by two different forms of skepticism. One community appears to be wary of accepting theories prematurely; the other community appears to be wary of dismissing theories prematurely.

[80] A concrete repercussion of these two forms of skepticism can be found in divergent attitudes towards the language of scholarly reporting.

Open Accounts versus Closed Explanations

[81] Scientists are apt to take issue with the idea that traditional humanities scholars are more likely to give interesting hypotheses or interpretations the benefit of the doubt. A scientist might well point out that many traditional humanities scholars are often skeptical of scientific hypotheses for which a considerable volume of supporting evidence exists. How, it might be asked, can a humanities scholar give credence to Freud's notion of the Oedipal complex while entertaining doubts about the veracity of Darwin's theory of evolution? I think there are two answers to this question -- one answer is substantial, while the second answer arises from an understandable misconception.

[82] The substantial answer has to do with whether a given hypothesis tends to preclude other possible hypotheses. The Oedipal complex might be true without significantly precluding other ideas or theories concerning human nature and human interaction. However, if the theory of evolution is true, then a large number of alternative hypotheses must be discarded. It is not necessarily the case that the humanities scholar holds a double standard when evaluating scientific hypotheses. If a scholar is motivated by theory-conserving skepticism (that is, avoiding false-negative claims), then a distinction must be made between those theories that claim to usurp all others, and those theories that can co-exist with other theories. The theory-conserving skeptic may cogently choose to hold a given hypothesis to a higher standard of evidence precisely because it precludes such a wealth of alternative interpretations.

[83] In the humanities, young scholars are constantly advised to draw conclusions that "open outwards" and to "avoid closure." This advice contrasts starkly with the advice given to young scientists who are taught that "good research distinguishes between competing hypotheses." From the point of view of the false-negative skeptic, a "closed" explanation greatly increases the likelihood of false-negative errors for the myriad of alternative hypotheses.

[84] This fear is particularly warranted whenever the volume of available data is small, as is often the case in humanities disciplines. A low volume of evidence means that no single hypothesis can be expected to triumph over the alternatives, and so claims of explanatory closure in data-poor fields are likely to be unfounded. For this reason, many humanities scholars regard explanatory "closure" as a provocation -- a political act intended to usurp all other views.

[85] Of course many scientific theories do indeed achieve a level of evidence that warrants broad acceptance and rejection of the alternative theories. Still, not all humanities scholars will be convinced that the alternative accounts must be rejected. I suspect that all researchers (both humanities scholars and scientists) tend to generalize from their own discipline-specific experiences when responding to work reported from other fields. Since humanities scholars often work in fields where evidence is scanty, the humanities scholar's experience shouts out that no knowledge claim warrants the kind of confidence commonly expressed by scientists. Objecting to scientific theories on this basis is clearly a fallacy, but it is understandable why scholars from data-poor disciplines would tend to respond skeptically to the cocky assurance of others. We will return to consider the issue of explanatory closure later, when we discuss Ockham's razor and the issue of reductionism.

[86] Having proposed this association between theory-discarding skepticism and science (on the one hand) and theory-conserving skepticism and the humanities (on the other hand), let me now retract and refine it. I do not think that there is any *necessary* association. The origin of this tendency, I propose, has nothing to do with the nature of scientific as opposed to humanities scholarship. I should also hasten to add that I do not believe that individual scholars are solely theory-discarding or theory-conserving skeptics. People have pretty good intuitions when to approach a phenomenon as a false-positive skeptic and when to approach a phenomenon as a false-negative skeptic.

[87] If there is no necessary connection between theory-discarding skepticism and science, and theory-conserving skepticism and the humanities, where does this apparent association come from? I think there are two factors that have contributed to these differing methodological dispositions. As already suggested, one factor relates to the quantity of available evidence or data for investigating hypotheses or theories. A second factor pertains to the moral and esthetic repercussions of the hypotheses. These two factors are interrelated so it is difficult to discuss each factor in isolation. Nevertheless, in the ensuing discussion, I will attempt to discuss each issue independently.

High Risk versus Low Risk Theories

[88] For the casual reader, one of the most distinctive features of published scientific research are those strings of funny Greek letters and numbers that often pepper the prose. Some statement is made, such as "X is bigger than Y," and this is followed in parentheses by something like the following:

X^2=8.32; df=4; p<0.02

There is some skill involved in understanding these numbers, but the essential message is conveyed by the value of p.

[89] In statistical inference, the value p is a calculated value that estimates the probability of making a false-positive error. If the researcher is endeavoring to avoid making a false positive claim, then the value of p should be as small as possible. As we have seen, depending on the circumstances, the researcher may wish to minimize the possibility of making a false negative error (i.e. theory-conserving skeptic). How does a researcher know

what type of error to minimize? Should the researcher be skeptical of negative claims or skeptical of positive claims? Should the researcher aim to conserve theories or discard them?

[90] The answer to this question is that it depends upon the moral (and esthetic) consequences of making one kind of error versus another kind of error. Consider, for example, the difference between civil and criminal cases in jurisprudence. Civil cases (such as trespassing) require comparatively modest evidence in order to secure a conviction ("preponderance of evidence"). Criminal cases (such as murder) require much more convincing evidence ("beyond a reasonable doubt"). These different standards of evidence are warranted due to the different moral repercussions of making a false-positive error. Securing the conviction of an innocent person in a murder trial is a grave blunder compared to convicting an innocent person of trespassing.

[91] Fields of inquiry that carry significant risks (such as medicine, jurisprudence and public safety) ought to have high standards of confidence. If the field is data rich, it is especially important to collect a sufficient volume of evident so the researcher can assemble a convincing case. If the field is data poor (such as often happens in jurisprudence), then one must expect to make a lot of errors; the moral repercussions of a false-positive versus a false-negative error will determine whether the researcher should adopt a theory-conserving or theory-discarding skepticism. In criminal law, one can expect many failures to convict guilty people in order to minimize the number of wrongful convictions.

[92] In contrast with legal proceedings, most scholarly hypotheses have marginal moral or esthetic risk. For example, whether a theory of the origins of Romanesque architecture is true or false has little moral impact. However, risk is never entirely absent. Suppose that a musicologist found evidence suggesting that one composer had plagiarized a melody from another composer. If the claim of plagiarism was in fact false, then the first composer's reputation would be unjustly tarnished. If that composer were still living, then a false claim of plagiarism would be morally reprehensible.

[93] To the knowledgeable statistician there is nothing new in this discussion. Modern statisticians have always understood the reciprocal relationship between false positive and false negative errors, and have long recognized that whether a researcher endeavors to reduce one or the other depends entirely on the attendant risks of making either error. In most traditional arts and humanities scholarship, making a false positive claim rarely has onerous moral or esthetic repercussions. Conversely, false-negative claims have often been seen as reckless.

[94] Perhaps the best known theory-conserving argument is *Pascal's Wager*. Unconvinced by the many proofs offered for the existence of God, Pascal asked what would be lost if the proposition were true but our evidence scant? Pascal argued that the repercussions of making a false-negative error were simply too onerous. He chose to believe in God, not because the positive evidence was compelling, but because he thought that the moral risk associated with wrongly dismissing the hypothesis would require an extraordinary volume of contradicting evidence (Pascal, 1669).

[95] Historically, statistical tests have been used almost exclusively to minimize false-positive errors. It is the community of theory-discarding skeptics who have made the greatest use of statistics. I suspect that this historical association between the use of statistical inference and false-positive skepticism may account for much of the widespread suspicion of statistical arguments among arts and humanities scholars. Yet there is nothing in statistical inference *per se* that is contrary to the traditional arts/humanities scholar's penchant for false negative skepticism. As statisticians well know, common statistical procedures are equally adept at serving the theory-conserving skeptic.

[96] As noted earlier, the science/false-positive and humanities/false-negative association is changing. Contemporary medicine has become more cognizant of the dangers of prematurely discarding theories. Concurrently, many arts and humanities researchers are becoming more aware of the problems of theory-conserving skepticism. In the case of music, several hundred years of speculative theorizing has led to the promulgation of innumerable ideas -- many of which surely lack substance {8}. Until recently, there was little one could do about this. The scarcity of pertinent data in many humanities fields simply made it impossible to satisfy statistical criteria for minimizing false positive errors. The opportunities to address these problems have

been immensely expanded due to the growing availability of computer databases, comprehensive reference tools, and the growing use of experiment-based data collection. We will return to these issues shortly.

Historical Fields

[97] Fields can be characterized according to whether the principal evidence or data arise from the past or from the future. Historical fields are fields whose fundamental data already exist. Archeology, paleontology and art history are examples of historical fields. In each of these fields, the principal phenomena of study are ones that occurred in the past. These phenomena are accessible for study only through the tenuous traces of currently existing data. Historical data might include paper documents, physical objects, oral histories, or unspoken memories. Normally, the existing evidence constitutes a proper subset of all of the pertinent evidence, most of which has been destroyed by the passage of time.

[98] It would be wrong to think of historical fields as principally belonging to the humanities. The sciences of astronomy, geology, and paleoanthropology are predominantly historical fields. Each of these sciences is concerned primarily with evidence of past events. Indeed, the preeminent historical discipline, it might be argued, is *astronomy*: the light that reaches astronomers' telescopes is typically hundreds or millions of years old. It is rare that astronomers get to study "current events."

Retrospective versus Prospective Data

[99] Historical data should not be confused with what may be called *retrospective* evidence or data. Retrospective data is evidence that is already in-hand -- evidence that is *known* to the researcher. *Prospective* data, by contrast, is data that is not yet available to the researcher. Prospective data includes evidence that will be collected in the future, but prospective data also includes existing evidence that a researcher has not yet seen -- such as data published in a forgotten article, or manuscripts in an overlooked archive.

[100] Note that prospective data can be entirely historical. Consider, by way of example, weather forecasting. We normally think of meteorologists testing their models by forecasting future weather, such as predicting tomorrow's weather, the weather next week, or the weather next year. However, most meteorological theories are tested using historical data. Given the antecedent data, a theory might be used to predict the weather on, say, March 2nd, 1972.

[101] Similarly, suppose that an ethnomusicologist formulates a theory based on a study of three hunter-gatherer societies. For example, the ethnomusicologist might theorize that matrilinear hunter-gatherers employ predominantly ascending melodic contours whereas patrilinear hunter-gatherers exhibit predominantly descending melodic contours. This theory might be tested by predicting specific cultural patterns in other hunter-gatherer groups. We might test the ethnomusicologist's predictions by carrying out new field research in as-yet-unstudied cultures. However, we could also test the ethnomusicologist's predictions against already existing data about other societies, provided the data is prospective rather than retrospective. Similarly, historians might test specific theories by predicting the contents of newly discovered (yet unopened) documents pertaining to a particular historical event.

[102] Of course in some areas of research, all of the pertinent data is already available. No amount of money will necessarily increase the volume of documents relating directly to Petrarch's life. In other words, all of the data is retrospective and researchers hold little hope of future prospective data. The loss of opportunities for prospective data removes the possibility of evaluating a theory by testing predictions. This situation has onerous repercussions for the affected area of research.

Pre-Data Theory and Post-Data Theory

[103] One of the most pernicious problems plaguing historical disciplines is the tendency to use a single data set both to generate the theory and to support the theory. Formally, if observation O is used to formulate theory T, then O cannot be construed as a predicted outcome of T. That is, observation O in no way supports T.

[104] The origin of the Theory of Continental Drift arose from observing the suspicious visual fit between the east coasts of the American continents and the west coasts of Europe and Africa. The bulge of north-west Africa appears to fit like a piece of a jig-saw puzzle into the Caribbean gulf. This observation was ridiculed as childish nonsense by geologists in the first part of the twentieth century. Geologists were right to dismiss the similarity of the coast-lines as *evidence* in support of the theory of continental drift, since this similarity was the origin of the theory in the first place. Plate tectonics gained credence only when independent evidence was gathered consistent with the spreading of the Atlantic sea-bed.

[105] Such "*post hoc* theorizing has particularly plagued evolutionary theorizing (see Gould, 1978; Gould & Lewontin, 1979; Lewontin, 1991; Rosen, 1982). Nevertheless, in some cases, evolutionary theories can arise that make predictions about yet-to-be-gathered data (such as the Trivers-Willard hypothesis). Good theories are *a priori;* that is, the theory suggests or predicts certain facts or phenomena before those facts are ascertained or observed.

[106] Fields that rely exclusively on retrospective data are susceptible to *post hoc* theorizing where hypotheses are easy to form and difficult to test. This is a problem that is endemic to many fields, especially historical fields (including astronomy). Nevertheless, careful attention to the underlying logic of a theory may permit testing of unexpected predictions of pre-existing prospective data. The fields of astronomy and evolutionary biology have demonstrated that there are many more opportunities for testing historical theories than is recognized by historians working in humanities disciplines.

Experimental versus Correlational Data

[107] A further distinction can be made between two types of prospective data. When making predictions about prospective data, a distinction can be made between phenomena that can be influenced by the researcher and phenomena that are beyond the researcher's influence. In some cases (such as weather forecasting), researchers have little or no opportunity to manipulate the initial conditions and observe the consequences. In other cases, researchers can initiate phenomena themselves or contrive or influence the initial conditions or context for some phenomenon, and then observe the ensuing consequences.

[108] Disciplines that can or cannot influence the phenomena under study are methodologically distinct. When significant interaction with the phenomenon is possible, scholars can carry out formal experiments. For example, a psychomusicologist can directly manipulate the timbre of a sound and determine whether listeners from different cultures perceive the sound as "more cute" or "less cute." By manipulating single variables, an experiment allows the researcher to infer causality. A properly designed experiment allows the researcher to demonstrate that *A* has affected *B* rather than *B* affecting *A*. By contrast, researchers in historical disciplines cannot carry out controlled experiments. There is no way to go back into the past to change a single variable, nor is there any way to construct an independent world and observe the effects of specific manipulations. In the language of empirical methodology, historical disciplines necessarily rely on *correlational* rather than *experimental* methods.

[109] In correlational studies, the researcher can demonstrate that there is a relationship or association between two variables or events. But there is no way to determine whether *A* causes *B* or *B* causes *A*. Moreover, the researcher cannot dismiss the possibility that *A* and *B* are not causally connected. It may be the case that both *A* and *B* are caused by an independent third variable. By way of illustration we might note that there is a strong correlation between consumption of ice cream and death by drowning. Whenever ice cream consumption increases there is a concomitant increase in drowning deaths (and vice versa). Of course the likely reason for this correlation is that warm summer days lead people to go swimming and also leads to greater ice cream consumption. In historical disciplines, one can never know whether the association of two events is causal, accidental, or the effect of a third (unidentified) event or factor.

Data Rich and Data Poor

[110] Of all the taxonomic distinctions made in this article, probably the most seminal is the distinction between *data-rich* and *data-poor* areas of research. Although the term "data" unfortunately implies something scientific, I intend the term to be construed in the broadest possible sense, meaning any information, observation, artifact, or evidence that may be pertinent to some theory, hypothesis, interpretation, or intuition. (In Latin, *datum*: a thing known, or passed around.)

[111] Data-rich disciplines are in principal able to uncover or assemble as much information, evidence, observations, etc. as they wish, limited only by financial resources. Data-poor disciplines have little control over the volume of pertinent data. As noted earlier, no amount of money will necessarily increase the volume of documents relating directly to a historical figure's life.

[112] There are four ways a field can be data-poor. One way is that the phenomenon itself is comparatively rare. It is difficult to study phenomena such as ball lightning, monosyllabic vowel-consonant verbs, white Bengali tigers, or multiple personality disorder. Few historical musicologists will experience the thrill of discovering a manuscript for an unknown work by a major composer.

[113] A second way by which a field may be data-poor, is that the data may be volatile or is quickly destroyed. For the paleontologist, soft body tissues disappear in a matter of years and so are difficult to study from fossilized rock samples. Some sub-atomic particles exist for less than a millionth of a second. For the psychomusicologist, the moment-by-moment expectations of a music listener are ephemeral and evanescent.

[114] A field may also be data-poor because the data is inaccessible. Archeological data is smothered by dirt. Neutrinos are thought to be everywhere in large quantities, but they have no electrical charge and no mass, so they resist interacting with any detection device. Although hundreds of thousands of amateur sound recordings are made each year, musicologists find them difficult to study: how does one assemble the recordings of Bach keyboard works performed by amateurs in 1999?

[115] Finally, data can simply be lost. The destruction of the famed ancient library at Alexandria transformed pre-Socratic philosophy into a notoriously data-poor field. A modern translation of *all* of the surviving pre-Socratic Greek texts runs to just 162 pages (Fitt, 1959). This includes the complete extant texts from the writings of Pythagoras, Thales, Anaximander and dozens of other classical thinkers. Musical examples abound: for example, not a trace remains of Dufay's *Requiem*.

Positivist Fallacy

[116] Data poor fields raise some special methodological concerns. One of these is the problem known as the *positivist fallacy.* If a phenomenon leaves no trail of evidence, then there is nothing to study. We may even be tempted to conclude that nothing has happened. In other words, the positivist fallacy is the misconception that *absence of evidence* may be interpreted as *evidence of absence.*

[117] Positivism had a marked impact on mid-twentieth century American psychology. In particular, the influence of logical positivism was notable in the behaviorists such as J.B. Watson and B.F. Skinner. The classic example of the positivist fallacy was the penchant of behaviorists to dismiss unobservable mental states as non-existent. For example, because "consciousness" could not be observed, for the positivist it must be regarded as an occult or fictional quality with no truth status (Ayer, 1936).

[118] If it is true that the positivist fallacy tends to arise from data-poor conditions, then it should be possible to observe this same misconception in humanities scholarship -- whenever data is limited. Consider, by way of example, the following argument from the distinguished historical musicologist, Albert Seay. At the beginning of his otherwise fine book on medieval music, Seay provides the following rationale for focusing predominantly on sacred music in preference to secular music:

> "Although much music did exist for secular purposes and many musicians satisfied the needs of secular audiences, the Church and its musical opportunities remained the central preoccupation. No better evidence of this emphasis on the religious can be seen than in the relative scarcity of both information and primary source materials for secular music as compared to those for the sacred." (Seay, 1975, p.2)

In other words, Seay is arguing that, with regard to secular medieval music-making, absence of evidence is evidence of absence. Since secular activities generated little documentation, we have almost no idea of the extent and day-to-day pertinence of medieval secular music-making. For illiterate peasants, "do-it-yourself" folk music may have shaped daily musical experience far more than has been supposed. Of course Seay may be entirely right about the relative unimportance of secular music-making, but in basing his argument on the absence of data, he is in the company of the most rabid logical positivist. The positivist fallacy is commonly regarded as a symptom of scientific excess. However, it knows no disciplinary boundaries; it tends to appear whenever pertinent data are scarce.

Parsimony versus Pluralism

[119] An important intellectual precursor of logical positivism can be found *Ockham's Razor.* William of Ockham promoted the idea that the number of factors entailed by an explanation should not be multiplied beyond those necessary. Modern philosophers more commonly refer to this as the *principle of parsimony* -- namely, that one should prefer the simplest hypothesis that can account for the observed evidence. Unessential concepts, factors, or causes should be excised.

[120] Of course the simplest explanation may not be the correct explanation. Biologists in particular have discovered that physiological processes are typically much more convoluted than would seem to be necessary. Nevertheless, there is methodological merit in eschewing unnecessary complexity. Every time an additional parameter or factor is introduced, the capacity for false-positive errors is increased considerably.

[121] By way of illustration, consider a hypothetical music theory that purports to explain every possible 8-note melodic phrase constructed using pitches within the range of an octave. (There are over 800 million possible phrases of this sort.) Mathematically, every conceivable 8-note pitch sequence can be perfectly modeled using just 7 parameters. Any music theorist can easily posit 7 plausible factors that influence the shape of a phrase. For example, a phrase might be influenced by scale type, contour shape, degree of chromaticism, Schenkerian line, pitch proximity, gap-fill tendency, stylistic period, etc. However, if a researcher claims to have a melodic model that accounts for all possible 8-note pitch sequences using just 7 factors, then the researcher has done no better than chance. Limiting the number of parameters or factors dramatically decreases the likelihood of constructing a spurious model or explanation.

[122] For the false-positive skeptic, the principal of parsimony holds merit, not because it reduces complex phenomena to simple phenomena, but because decreasing the number of variables reduces the chances of making a false-positive error. While increasing the number of contributing factors can make a model more realistic, regrettably, it also greatly increases the capacity for self-deception.

Three Faces of Reductionism

[123] There are at least three ways of interpreting the term *reductionism.* One is the methodological injunction to use the least number of variables possible when formulating a theory. This view of reductionism is synonymous with the principal of parsimony, which we have just discussed. A second way of understanding reductionism is the "divide and conquer" method of research. A third interpretation of reductionism is the "nothing but" mode of explanation. These latter two notions of reductionism are described below.

[124] "Divide and conquer" reductionism endeavors to elucidate complex phenomena by isolating constituent relationships. Classically, the principal research tool for this form of reductionism is the concept of "control." It

is commonly thought that control entails holding one or more factors constant while the "independent variable" is manipulated and the "dependent variable" is observed. However, control more commonly entails randomizing the potentially confounding variables. In taking a political poll, for example, pollsters expect that the number of variables influencing a particular opinion is very large. It is hopeless to assume that one can hold constant such a large number of factors. Consequently, researchers seek a random sample with the hope that unknown influences will tend to cancel each other out. The formal statistical argument in support of random sampling is quite compelling, so there is considerable merit to this method of control.

[125] Using such methods of control, it becomes possible for a researcher to investigate the effect of a given factor on some complex phenomenon. By investigating one factor at a time, it is often possible to build a sophisticated model or theory of the phenomenon in question. When the number of factors is more than five or six, the divide and conquer strategy often becomes intractable due to the explosion of possible interactions between purported factors. Nevertheless, the approach can still help identify important relationships in real-world phenomena.

[126] The more contentious form of reductionism may be called the "nothing but" mode of explanation. A reductionist attempts to explain complex phenomena as merely the interaction of simpler underlying phenomena; explanation proceeds by accounting for complex wholes in terms of simpler components. In this form of reductionism, the researcher aims to make statements of the form "*X is nothing but Y*."

[127] Used in this sense, reductionism can be contrasted with what is sometimes called *holism*. A `holist' expects to explain phenomena as being greater than the sum of its parts (a process dubbed *synergism* by Buckminster Fuller). Frequently, synergism leads to "emergent properties" where complex phenomena cannot be predicted even when a thorough understanding exists of the underlying constituent phenomena.

[128] In contrast to the holist, the "nothing but" reductionist seeks to explain all complex phenomena as convoluted manifestations of a handful of fundamental causes or interactions. Culture is just sociology, sociology is just psychology, psychology is just biology, biology is just chemistry, and chemistry is just physics.

[129] One cannot help but be impressed by the breathless grandiosity of this program. If such a scientific reductive synthesis is true, it will represent one of the pinnacle achievements of human inquiry. If it is false, it will represent one of the preeminent intellectual blunders in human history.

[130] Humanities scholars of many stripes have derided the reductionist project. Much of the objection originates in the unsavory esthetic repercussions of `nothing but' reductionism. It is argued that such reductionistic accounts "explain" only in the sense of *making flat* (*ex planum*). The world as an enchanting place is transformed into a prosaic, colorless, and seemingly senseless enterprise. Among humanities scholars, musicians and musicologists have been among the most vocal critics of `nothing but' reductionism. Music theorists explicitly embrace complexity and scorn simplicity. {9}. John Cage cautioned strongly against such "logical minimizations." Moreover, Cage was prescient in recognizing that this reductive tendency is not limited to the sciences. It is surprising where one can find such "nothing but" forms of reductionism.

[131] Consider, once again, postmodernism. The postmodernist/deconstructionist philosophy advocates the unpacking of concepts and utterances in terms of socially constructed roles and power relations (e.g. Hacking, 1995). Postmodernism has helped to expose innumerable subtle and not-so-subtle ways in which ostensibly rational discourse manifests convoluted forms of dominance and control. But postmodernism goes much further. The most abstract principles of law, philosophy, and even science are best understood from the point-of-view of politics: *everything reduces to politics*. Notice that in this formulation, postmodernism and deconstruction bear all the hallmarks of *nothing-but* reductionism. Any thought you care to express can be reduced to a political motive. A sociobiologist may believe a social phenomenon to be ultimately reducible to underlying chemical interactions. But the postmodernist trumps this reductionism by viewing all scientific discourses as ultimately reducible to power ploys. As in the case of the scientific reductive synthesis, one cannot help but be impressed by the breathless grandiosity of such postmodernist patterns of explanations.

[132] There is, I would suggest a more helpful way of understanding the value of reductionism while avoiding some of the more unsavory excesses (in both the sciences and the humanities). A helpful distinction is to treat "reductionism" as *a potentially useful strategy for discovery* rather than *a belief about how the world is*. Concretely, the postmodernist might *use* the assumption of hegemony as a technique to help unravel a complex behavior. Similarly, the sociobiologist might *use* the assumption of a recessive gene as a technique to help analyze a personality trait. In both cases, there are dangers in assuming that the tool is the reality. But in both cases, there remains the possibility that the reductive explanatory principle proves useful in understanding the phenomenon in question.

Humanistic and Mechanistic Beliefs

[133] Our understanding of reductionism can be aided by contrasting the terms *reductionism* and *holism* with the philosophical distinction between *humanistic* and *mechanistic* views. The latter concepts might be defined as follows:

- **Humanistic:** A belief in spirit and consciousness as fundamental, and not reducible to mechanical descriptions.

- **Mechanistic:** A belief in a mechanical conception of life and consciousness. A belief that there is no essential mystery or enigma -- there is only our ignorance of how things work.

[134] Humanism and mechanism (as defined above) are beliefs, whereas reductionism and holism (as I've defined them) are methodological approaches. It is true that researchers who hold a mechanistic view of the world also tend to prefer reductionistic methods. It is also true that researchers who hold a humanistic view of the world tend to prefer or advocate holistic methods. However, there is no necessary link between humanism and holism, nor between mechanism and reductionism. There are many scientists (especially those working in the areas of complexity and chaos) who hold a mechanistic view of the world but who presume that complex interactions can lead to emergent properties that cannot be predicted (e.g., Anderson, 1972; Gell-Mann, 1994; Gleick, 1987; Pagels, 1988). In addition, a researcher can cogently hold a humanistic view of the origins of human behavior, yet rely on reductionism as a useful method for investigation. That is, one need not *believe* that human behavior is mechanistic in order to use reductionism as a way of probing the complexities of the world. Using reductionism as a research strategy does not commit a researcher to a mechanistic world-view. Similarly, analyzing a phenomenon as a holistic emergent property does not thereby transform the researcher into a spiritualist.

A Quantitative Role

[135] Earlier we noted that "empiricism" simply means knowledge gained through observation. For many critics of empiricism, it not the idea of observational knowledge *per se* that raises concerns, but empiricism's widespread reliance on quantitative methods.

[136] Perhaps the preeminent concern is that quantitative methods force phenomena into numerical categories that may or may not be appropriate. A researcher, for example, might ask listeners to rate musical excerpts on a scale from 1 to 10, where 1 represents "maximum sadness" and 10 represents "maximum happiness." This practice is open to innumerable objections: happiness and sadness may be independent phenomena that do not exist on some unified continuum; the musical excerpt may not retain a consistent character throughout the passage; a "poignant" passage might be both "happy" and "sad" simultaneously; a passage might be recognizable as *intending* to portray happiness, but a listener may find the portrayal unconvincing, and so "sadly" a failure; the numerical judgments may be uninterpretable (is the value 2 intended to be half as sad as the value 1?), etc.

[137] Concerns such as these actually form much of the fundamental curriculum for training in quantitative methodology. For example, empiricists are taught that any judgment scale should use a single adjective (ranging from "least *X*" to "most *X*") rather than using mixed adjectives ("most *X*" to "most *Y*"). Similarly, empiricists

learn that measurements are never to be construed as direct indices of actual phenomena, and operational definitions should not be reified. Statisticians have devised completely independent analytic procedures, depending on the properties of various measurement scales.

[138] For many humanistically-inclined scholars, however, there remains something inherently wrong about quantifying human experiences -- especially those experiences related to human attachment, esthetic experience, and spiritual life. Many scholars would agree with Renato Poggioli's view that the technical and quantitative have their place, but not in the arts:

> "Technicism" means that the technical genius invades spiritual realms where technique has no raison d'être. ... It is not against the technical or the machine that the spirit justly revolts; it is against this reduction of nonmaterial values to the brute categories of the mechanical and technical." [p.138]

Once again, let me respond to this view by distinguishing *methodologies of scholarly inquiry* from *philosophical beliefs about the nature of the world.* Lest this distinction seem too abstract, consider the following extended illustration, which draws a parallel to scholarly attitudes regarding the use of writing and musical notation.

[139] Socrates famously criticized the new fangled invention of writing. He rightly pointed to a number of predictable, yet questionable, consequences of relying on written texts. Specifically, Socrates predicted a decline in the importance of rote memory, and the waning of oratory skills.

[140] Socrates' predictions have been amply proved correct. Few modern children can recite more than a single poem, politicians rely on teleprompters, and humanities scholars make public presentations with their heads buried in dense texts that leave listeners confused. Socrates' legitimate criticisms notwithstanding, writing caught on. In fact, writing was soon recognized as providing an invaluable window on previously unknown phenomena. With writing, for example, the Greeks discovered grammar. By removing speech from the ephemeral moment, the ancients discovered "parts of speech" (nouns, adjectives, particles, etc.) as well as tenses, conjugations, sentences, plots, and other structures. In short, the invention of writing provided an unprecedented opportunity to better understand language, and (paradoxically) speech.

[141] An almost identical history attended the advent of musical notation. Music theorizing was common long before music was written down. But music notation unquestionably inspired and facilitated the growth of music theory in the West. As in the case of written language, musical notation allowed those who study music to identify patterns of organization that would otherwise be difficult or impossible to discern.

[142] Of course, like Socrates, musical notation has drawn its critics. Jazz musicians are likely to resonate with the observations of a nineteenth century Arab traveler to Europe, Faris al-Shidyaq:

> "The Franks [Europeans] have no `free' music unbound by those graphic signs of theirs ... so that if you suggest to one of them that he should sing a couple of lines extempore ... he cannot do so. This is strange considering their excellence in this art, for singing in this fashion is natural and was in use among them before these graphic signs and symbols came into being." [As quoted in Nettl, 1985, p.123]

[143] A perhaps unfortunate repercussion of musical notation has been the reification of notation *as* music. The very noun "music" has today acquired meanings that would have confounded ancient musicians. In modern times it is possible for "music" to fall off a stand or to be eaten by one's dog. Consider philosopher Nelson Goodman's well-known conception of the identity of the musical work:

> "A score, whether or not ever used as a guide for a performance, has as a primary function the authoritative identification of a work from performance to performance. Often scores and notations -- and pseudo-scores and pseudo-notations -- have such other more exciting functions as facilitating transposition, comprehension, or even composition; but every score, as a score, has the logically prior office of identifying a work." (Goodman, 1976/1981; p.128).

For Goodman, the notion of the existence of a musical work devoid of any score is a highly complex and thorny philosophical issue. In Goodman's view, the very identity of "music" is intimately linked and equated with material notational artifacts of a certain sort. This is what is meant by "reification."

[144] As in the case of written language and musical notation, quantitative methods provide (1) important opportunities for glimpsing otherwise invisible patterns of organizations, and (2) similar opportunities for reification and fetishism. Scholarly attitudes toward musical notation are rightly mixed: notation has provided extraordinary opportunities for scholarly inquiry, but it has also spawned some moot and questionable beliefs regarding the nature of the musical world.

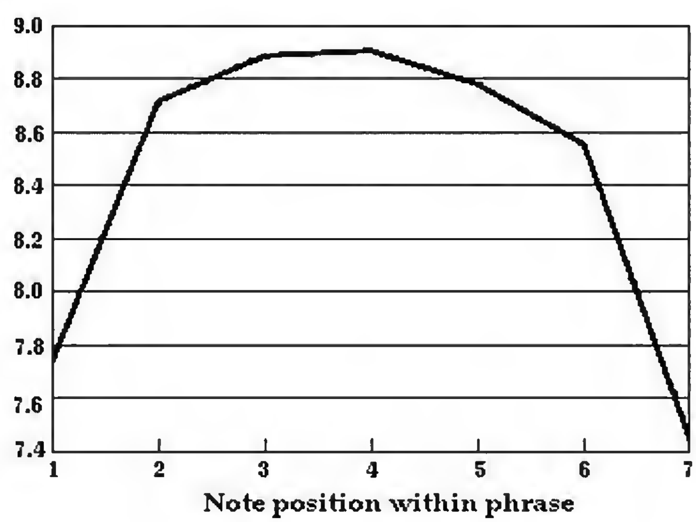
[145] In the case of applying quantitative methods in music scholarship, we are a long way away from such excesses. On the contrary, music scholarship has barely begun to take advantage of the genuine opportunities provided for better understanding musical organization. Of the many examples that can be used to illustrate the promise of quantitative empirical methods, two examples must suffice. My first example relates to the concept of the "melodic arch" whereas the second example relates to the concept of "gap fill".

The Melodic Arch

[146] For centuries, music theorists have drawn attention to the so-called "melodic arch" -- a presumed general tendency for melodic phrases to ascend and then descend. An example of an arch-shaped phrase might be the opening phrase of *My Bonnie Lies Over the Ocean*. Unfortunately, there are also lots of counter-examples: *Joy to the World* and the *Star Spangled Banner* are just two of many melodies that exhibit "convex" initial phrases.

[147] What is one to make of the concept of the "melodic arch"? Is it true that there is a general arch tendency in musical phrases? Or have textbook writers simply been selective in their examples?

[148] Huron (1996) carried out a study involving more than 36,000 melodic phrases sampled from European folksongs. The first question to resolve is one of definition: what is an "arch"? One way to define an arch is that all of the notes in the first half of the phrase rise upward in pitch, while all of the notes in the second half of the phrase move downward. A less restrictive definition might simply require that the average pitch-height of the initial and final notes of a phrase are lower than the average pitch-heights of the mid-phrase notes. Alternatively, one might determine phrase contours only after non-structural tones have been discarded. Without resolving the issue of what we mean by an "arch", Huron's study used several different operational definitions and found that the results were the same no matter how one defines an arch. By way of illustration, Figure 1 (below) shows the results of just one way of addressing the matter. The figure shows what happens when the pitch heights of 6,364 seven-note phrases are averaged together.



[149] In Huron's study, each of the alternative notions of a "melodic arch" converged on the same answer. Although there are many individual phrases that do not exhibit an arch-shape, the great majority of phrases do indeed have a roughly ascending-descending contour. That is, the results are consistent with a general theoretical notion of a melodic arch (at least in Western folksong melodies). One might suppose that averaging together thousands of melodic phrases constitutes the epitome of quantitative lunacy. Yet, such simple quantitative procedures can prove remarkably useful in addressing certain kinds of musical questions.

Gap Fill

[150] A common criticism of empirical studies in music is that they merely confirm our intuitions. A good counter-example is provided by the phenomenon of "gap fill." For 500 years, music scholars have observed that large melodic leaps tend to be followed by changes of melodic direction. This phenomena goes under a number of names, but let us use Leonard Meyer's terminology: "gap fill."

[151] In a series of empirical studies, Paul von Hippel (2000a, 2000b, von Hippel & Huron, 2000) carried out extensive empirical investigations of the gap fill concept. The results are not at all consistent with music theorists' intuitions about gap fill. The story has two parts:

> (1) It is indeed the case that the majority of large intervals tend to be followed by a change in melodic direction. This pattern occurs in melodies from cultures spanning five continents and 500 years. This pattern is evident both for immediate pitch continuations, as well as delayed pitch continuations.
>
> However ...
>
> (2) If you completely scramble the order of notes within a melody, you end up with "random" melodies that tend to have exactly the same amount of gap fill as the original melodies themselves. This pattern occurs in melodies from cultures spanning five continents and 500 years.

[152] The fact that scrambled (randomly reordered) versions of the same melodies produce the same gap fill tendency suggests that gap fill is not a consequence of compositional intent.

[153] There is a straightforward explanation for why this happens -- a phenomenon that statisticians call "regression toward the mean". A large leap will have a tendency to take the melody towards the upper or lower extremes of a melody's range. Having landed (say) near the top of the range, the melody has little choice but to continue with one of the lower notes. In real music, the closer the leap is to the extremes of the range, the more likely the contour changes direction. When a leap lands in the middle of the tessitura, reversing direction is no more common than continuing in the same direction.

[154] Quantitatively, this account is very strong. After accounting for regression-toward-the-mean, there is **no** residual melodic behavior that can be attributed to a hypothetical principle of gap fill. While research on peripheral aspects of this issue continues, at this point it appears that "gap fill" is a musical concept without any reality in a large and diverse sample of actual notated music {10}.

The Promise of Quantitative Methods

[155] As I have argued, quantitative methods are important for the same reason that musical notation can be important: like musical notation, quantitative methods allow us to observe patterns of organization that might otherwise be difficult or impossible to decipher. For the new empiricist, an interest in quantitative methods has nothing to do with science. It has everything to do with becoming a more observant music scholar.

[156] Consider, finally, the value of quantitative methodology in resolving how assertions are made in humanities scholarship. To the outsider, it often appears that the essence of scholarly debate is that one scholar

believes that *X* is true, whereas another scholar believes that *X* is false. Most scholarly disagreements, however, relate to subtle shades of certainty. Consider, for example, the following assertions:

1. Tchaikovsky most certainly did not commit suicide.
2. Tchaikovsky very likely did not commit suicide.
3. Tchaikovsky probably did not commit suicide.
4. Tchaikovsky perhaps did not commit suicide.
5. Tchaikovsky may or may not have committed suicide.
6. Tchaikovsky perhaps committed suicide.
7. Tchaikovsky probably committed suicide.
8. Tchaikovsky very likely committed suicide.
9. Tchaikovsky most certainly committed suicide.

Most Tchaikovsky scholars suspect that Tchaikovsy did not commit suicide, but they disagree about the strength of the evidence, and hence they disagree about how a scholar should express this idea. Different scholars will accept (2), (3), or (4) above, but (1) will be considered excessive. While a given scholar may write (2) in a peer-reviewed journal, the ire of his or her colleagues may be provoked if his/her ensuing book prints (1) instead. Such are the proper nuances of scholarship.

[157] Scholars familiar with quantitative methodology will immediately recognize that the disagreement amounts to uncertainty about the value of p (described earlier) -- namely, the probability of making a false positive claim. In empirical research, the potential for mischief in reporting this idea would be circumvented by simply reporting the statistical confidence level.

[158] Quantitative methods provide little benefit when the amount of data is as miniscule as that pertaining to Tchaikovsky's death. But there are innumerable musical issues where quantitative methods are indispensable and powerful. Conductors may pride themselves on their unprejudiced golden ears, but economists Claudia Goldin and Cecilia Rouse have assembled the concrete numbers comparing blind and non-blind auditions: the results are consistent with a rampant and systematic discrimination against female orchestral musicians (Goldin & Rouse, 2000).

[159] In assessing the writings of another scholar, how are we to know whether the writer is guilty of using exaggerated rhetoric? Like the Tchaikovsky researcher commenting on Tchaikovsky's death, scholars may rightly wonder whether the assertion of, say, a feminist scholar, is being overstated or understated. But for those who understand quantitative methods, the numbers can be far more compelling -- and far more damning -- than any rhetorical flourish.

Using the Right Methodology

[160] By now it should be clear that I regard methodologies as tools for conducting research, not as philosophical belief systems. Like all tools, a given methodology is suitable for certain kinds of research but not other kinds. In pursuing a research program, thoughtful scholars will take stock of the conditions of their area of research and choose a methodological approach that best caters to the research goals -- in light of the opportunities and dangers inherent in specific tasks. The most appropriate methodology may change depending on the specific task or hypothesis being addressed.

[161] The principal impediment to careful selection of field-appropriate methods is the methodological inertia to be found in most disciplines. Researchers typically are taught only a single methodology and so expect this method to be applicable (to a greater or lesser degree) in virtually all research tasks. In the words of Abraham Maslow, to the person who holds a hammer, all the world looks forever like a nail.

[162] Even when researchers tend to use field-appropriate methods, we are often insensitive to the subtle changes in a field that ought to cause us to revisit and revise our methodological strategies and commitments. In

the remaining sections, we consider some of the misconceptions and failures that attend either (1) failing to recognize field-specific differences, or (2) failing to recognize changing conditions within a field of research.

Understanding Humanities Methods

[163] Scientists sometimes express dismay at the low levels of evidence that appear to be common in humanities disciplines. These views are often misplaced for two reasons. First, many humanities activities address `low risk' hypotheses in the sense that committing a false-positive error has modest repercussions. Second, data-poor disciplines simply cannot be expected to satisfy high standards of evidence.

[164] Faced with often paltry volumes of data, most scientists would never consider pursuing the sort of research projects found in the humanities. The scientist might be tempted to conclude that no knowledge claims ought to be made. *However, this presumes that false-negative errors have no moral repercussions.* It may be the case that `the lessons of history' are poorly supported and unreliable, but what are the consequences of concluding that it is impossible to learn from history? Historians are right, I believe, to try to make sense of incomplete and eclectic historical evidence -- since our failure to learn from this material may doom us to repeat past mistakes.

[165] In general, it should not be surprising that researchers in data-poor fields are typically oriented to theory-conserving skepticism rather than theory-discarding skepticism. When data is scarce, pursuing a theory-discarding skepticism means that one must always conclude that no conclusion is possible: no hypothesis can be supported, no theory is tenable. Any scholar having this disposition will naturally abandon the discipline.

[166] There are circumstances, however, where the dismay expressed by scientists concerning evidence in humanities disciplines is proper and appropriate. Specifically, these criticisms are warranted (1) when the risks of committing a false-positive error have significant moral (or esthetic) repercussions, and (2) when the field is not, or need not be, data poor. Both of these circumstances arise with some regularity in traditional humanities disciplines. Moreover, either one of these circumstances necessitates significant changes in methodology. An instructive historical example may be found in the splitting of the social sciences from the humanities.

The Social Sciences Split: Risky Hypotheses and Data Riches

[167] Humanities disciplines deal with human behavior, civil society, and culture. Humanities scholars regularly make claims about human nature, about moral and immoral conduct, and render advice about political, educational and cultural institutions. Scholars' views concerning these areas of human affairs can, and often do, have significant impact. At the end of the nineteenth century, the social sciences began to drift away from traditional humanities approaches precisely because thoughtful scholars recognized the need for higher standards of evidence in support of knowledge claims, especially those claims that might influence public attitudes and public policy.

[168] In recognizing the risks of committing false-positive errors, social scientists were right to initiate changes in their research methods. Contributing to this revolution in methodology was the realization that the social sciences could conduct research that would significantly increase the volume of evidence that could inform researchers' theorizing.

[169] Over the decades, a number of humanities scholars have criticized contemporary psychology and sociology for adopting methods more commonly associated with the physical sciences. However, these criticisms are based on the false assumption that disciplines are defined, not only by their subject matter, but also by their methods. As we have seen, methods arise not from the subject of research, but by the riskiness of the hypotheses, by the availability of pertinent data, by the ability of researchers to observe the effects of *a priori* manipulations, and by the opportunity to collect evidence independent from the original evidence used to formulate some theory or interpretation.

[170] It is wrong, I believe, to portray methodologies as competing philosophical allegiances. It is not a question of whether "scientific" methods prevail over interpretive, hermeneutic, phenomenological, or other traditional humanities methods, or vice versa. The question is whether researchers use the best methodology (or `basket' of methods) for the task at hand.

[171] To many scholars, it appears that over the course of the twentieth century, the humanities "lost" a number of disciplines -- including linguistics, archeology, psychology, and (to a lesser extent) anthropology and sociology. I disagree. The subject matter of these disciplines has changed little over the past century. Linguists are still interested in the origins, structures and acquisition of human languages. Archaeologists are still interested in how artifacts inform us about past human civilizations. Psychologists are still interested in human thoughts and motivations. Sociologists and anthropologists are still interested in the nature of human interaction and the nature of culture. In each discipline, human beings and human lives remain central. What has changed for these disciplines is primarily the volume of available evidence -- and consequently the opportunities to address more refined questions using methods that better exploit the expanded data resources.

[172] The prospect of gaining access to increased data is not merely an opportunity to be taken or ignored, as one pleases. Where pertinent data is readily available, it is morally reprehensible not to use it since failing to use the data increases the likelihood of making *both* false-positive and false-negative errors. In short, empirical data deserves our attention for precisely the same reason that small amounts of historical data warrant the historian's best interpretive efforts: failing to attempt to learn from the information at hand is to encourage and condone ignorance {11}.

Particle Physics: The Repercussions of Decreasing Data

[173] Although circumstances can open the flood-gates of data, circumstances can also close them. Admittedly, it is less common for a discipline to experience a reduction in the volume of data, but it does happen. The field of particle physics is arguably such a field. The very success of sub-atomic physics has pushed the frontier of study to more and more esoteric corners of reality. Particle physicists cannot carry out experiments without access to enormously costly machinery. After spending roughly $2 billion preparing to build the super-conducting super-collider (SSC), in 1993 the U.S. government decided to abandon the venture as too costly. Although particle physicists can continue to collect data, physicists have few opportunities to collect data that is pertinent to the latest theoretical models and issues.

[174] Even if the SSC had been built, its utility would have been limited. The most developed theories of physical reality exceed our abilities to test them. For example, in order to test hypotheses arising from *superstring theory,* it has been estimated that a suitable particle accelerator would need to be 1,000 light-years in circumference (Horgan, 1996; p.62). With the increasing scarcity of pertinent data, sub-atomic physics is slowly being transformed into a purely theoretical enterprise. Already, quantum physics has attracted innumerable competing interpretations with little hope that tests will ever be done that might prune away the incorrect interpretations. Nobel laureate, Sheldon Glashow expresses the malaise in his field as follows: "contemplation of superstrings may evolve into an activity ... conducted at schools of divinity by future equivalents of medieval theologians." (Glashow & Ginsparg, 1986; p.7).

[175] Glashow's allusion to theology is derisive. But particle physicists may need to get used to the apparently inevitable methodological transformation that awaits their discipline. Humanities scholars can be forgiven for shedding crocodile tears: for centuries, historians have had to struggle to make sense of manuscript fragments that they knew would never be made whole. When data is finite, interpretation is the only scholarly activity that remains. Moreover, the interpretive, hermeneutic enterprise is an activity that remains of value.

Musicology: The Repercussions of Increasing Data

[176] While sub-atomic physics is moving into a period of data scarcity, the reverse situation appears to be happening for music. As noted earlier, technical and organizational innovations can transform data-poor fields

into data-rich fields. Over the past 25 years, such innovations have arisen in many areas of musical study -- following the trends of such disciplines as linguistics, education and anthropology. Contemporary music scholars have access to computational and database resources, comprehensive reference tools, high quality data acquisition methods, sophisticated modeling techniques, and other innovations that make it far easier to collect, analyze and interpret musically-pertinent evidence and artifacts. There is hardly any area of music that cannot benefit from the increased resources, and from the ensuing opportunity to adopt more rigorous standards of evidence. This includes areas such as manuscript studies, poietics, history, iconography, analysis, performance, pedagogy, reception, esthetics and criticism, phenomenology, social and critical theory, cultural studies, cultural policy, media, and ethnology. Not all areas of music scholarship have, or will be touched by the expanding resources. Nor will speculative and creative music philosophy entirely lose its value.

[177] The changing landscape in musicology towards more empirical approaches is not a displacing of the humanities spirit by an antithetical scientific ethos. It is fundamentally a response to a clearer epistemological understanding of the role of methodology. Changing conditions simply allow us to be better music scholars, to embrace higher standards of evidence, and to be more acutely aware of the moral and esthetic repercussions of our knowledge claims, including claims that something is unknowable or that some phenomena ought not to be investigated. Our strongest criticisms should be levied at those who insist on speculative discourse when the resources are readily available to test such knowledge claims.

Impact Assessments in Humanities Discourse

[178] The above discussion has only cursorily addressed the issue of evaluating the moral and esthetic repercussions of various knowledge claims. Few aspects of humanities discourse are in greater need of discussion. I believe it is imperative that humanities scholars not be cavalier about the impact and importance of ideas. It is dangerous to suppose that, in comparison to *technologies* (with their considerable potential for mischief), *ideas* are somehow fragile and innocent. Karl Marx never failed to denigrate what he called "mere ideas." Philosophers, he said, have been content simply to talk about the world, with little interest in changing it. It is unfortunate that Marx never lived to see the cruel irony of his words. No other individual had so marked a *moral* effect on twentieth century lives as Karl Marx. Yet Marx himself was the quintessential closeted philosopher. Before letting an idea loose on the world, ideas ought to be subject to the same environmental impact assessments we apply to roadways and chemicals. Half-baked ideas have been just as disruptive and damaging as any technological innovation -- probably more so. It is important that humanities scholars stop underestimating our power to change the world. At the same time, it is important not to underestimate our culpability when we get things wrong.

Methodology as Pot-hole Guides

[179] Possibly the most pervasive misconception about methodology is that scholarly methods provide algorithms for carrying out research. According to this view, a methodology is a sort of recipe that scholars follow in the course of their studies. In this view, the function of epistemologists is presumed to be to concoct increasingly refined and more detailed methodological algorithms. The origin of this view may be linked to similar misconceptions about procedures in mathematical proofs. While the deductive procedures used by mathematicians are indeed rule-bound, mathematical research itself is a much more woolly-headed enterprise.

[180] As noted in the Part I, in the twentieth century, the idea of "methodology as algorithm" has come under sustained and devastating attack (Agassi, 1975; Feyerabend, 1975; Gellner, 1974; Kuhn, 1962; Laudan, 1977; Popper, 1934; Polanyi, 1962; Quine, 1953; and others). Many of these attacks have come from authors whose motivation was a defense of the rationality of science. The overwhelming conclusion from these critiques is that no known set of rules can guarantee the advance of knowledge. Moreover, as we have seen, even the most flexible known methodology `rule' yet proposed, Feyerabend's *anything goes,* fails to be born out by observation.

[181] Of the various efforts to reformulate our understanding of scholarly methodology, one of the best informed and most nuanced has been the view offered by the epistemologist Jagdish Hattiangadi. In his *Methodology without Methodological Rules,* Hattiangadi (1983) argues that, like scientific theories, methodological theories are activities of discovery, for which there are not fixed rules. The scholar who slavishly follows a fixed methodology will ultimately make an onerous mistake.

[182] Hattiangadi regards fields of scholarship as debating traditions that develop problems and criteria as they go. Although rationality is tradition-bound, rationality is not constrained solely by what we believe. What methodologists discover is a series of guidelines or heuristics.

[183] In our long history of making mistakes, scholars have come to identify common `pot-holes' on the road to understanding. Humanities scholars have learned to recognize and avoid a multitude of logical and rhetorical fallacies, including *ad hominem* arguments, appeals to authority (*ipse dixit*), the naturalist fallacy, the positivist fallacy, reification or hypostatization, and a host of pitfalls in forming historical explanations (Elster, 1989; Fischer, 1970; Roberts, 1996). Similarly, contemporary scientists have identified innumerable additional dangers. Among these dangers are the problem of hindsight reasoning, experimenter bias, ceiling effects, demand characteristics, the multiple tests problem, the third variable problem, cohort effects (Schaie, 1986), and the reactivity problem (Webb, Campbell, Schwartz, Sechrest, & Grove, 1981). These (and many other problems) are all well documented, and in many cases effective guidelines have been devised to recognize, avoid or minimize their detrimental effects on scholarship.

[184] Researchers are free to choose or develop their own methodology -- whether deductive, empirical, phenomenological, or whatever. But the pursuit of knowledge is best served when scholars learn from the various existing debating traditions. Although there is no detailed road-map for pursuing research, there exist sketches of well-documented pot-holes that others scholars have already encountered. It is important for scholars to be aware of these known hazards and for disciplines to keep abreast of methodological discoveries. Methodology is not simply some abstract specialty of philosophy. It is a utilitarian cross-disciplinary consultancy that offers pragmatic day-to-day assistance for all researchers.

[185] Here, regrettably, postmodernism has done humanities scholarship a grave disservice. Many otherwise thoughtful people are convinced there is no possibility of rigor, and that methodology is a dangerous illusion. As a result, an entire generation of students in the arts and humanities has been deprived of adequate practical education relating to methodology. To the postmodernist skeptic, one must respond with the reverse skepticism: What if there are truths? What if some truths are knowable? What if some interpretations are better than others? What if we fail to learn from the evidence that is available to us?

Conclusion

[186] By way of review, the basic arguments I have presented can be reconstructed and summarized as follows:

1. Postmodernists are right to note that knowledge claims do not take place in a moral vacuum. Theories, hypotheses, interpretations and opinions carry moral (and esthetic) repercussions. Moreover, choosing to avoid making knowledge claims is similarly an act with moral consequences.
2. Anyone wishing to make any knowledge claim about the world, has no choice but to navigate the treacherous path between false positive and false negative errors. This includes claims that say `I don't know' and `We cannot know.' There is nothing epistemologically safer about these negative claims compared with the corresponding positive claims `I know' or `In principal, we can know.'
3. The "Problem of Induction" is intractable and omnipresent: no amount of observation can establish the truth of some proposition. This problem applies not only to empiricism, but also to the critiques of empiricism offered by anti-foundationalist writers like Feyerabend. No amount of observation about the history of science can establish the general claim that the enterprise of science is irrational or arational.
4. Despite the problem of induction, observation remains indispensable to knowledge in ways we do not understand. Our very biological machinery has evolved to facilitate acquiring knowledge about the world.

We can show that observations are consistent with some theories and not other theories -- even though we cannot *prove* that one theory is better than another.

5. Fields of study differ according to the volume and quality of available evidence ("data") used to support or assess different claims, views, interpretations, or theories.
6. When data are inaccessible or non-existent, the field is susceptible to the *positivist fallacy* -- that absence of evidence can be interpreted as evidence of absence.
7. Data-poor fields are unable to support research whose goal is to minimize false-positive claims. Theory-discarding skeptics therefore avoid pursuing research in data-poor fields; they conclude that no conclusions can be drawn from the available data.
8. Other scholars will recognize the possibly onerous moral repercussions from failing to attempt to learn from small amounts of data/evidence. Data-poor fields will attract only theory-conserving skeptics, that is, scholars whose goal is to minimize false-negative claims.
9. When the volume of data is small, false-negative skeptics are logically consistent when they support multiple alternative hypotheses or interpretations. Pluralism is therefore preferred over parsimony. Conclusions are open rather than closed.
10. Unfortunately, scholars working in data-poor fields will typically make innumerable false-positive errors. That is, many ideas will be promulgated that lack merit.
11. Data-rich fields provide greater power for hypothesis testing. More stringent criteria allow testing that minimizes false-positive claims. As a result, competing hypotheses can be rejected with some assurance. Parsimony is therefore preferred to pluralism. Researchers aim for closed explanations.
12. Data can also be characterized as retrospective or prospective. Retrospective data invites two methodological problems. First, retrospective data is susceptible to unfettered "story-telling:" scholars are adept at formulating theories that account for any existing set of data. That is, it is tempting to use retrospective data both to formulate an explanatory theory and to provide evidence in support of the theory. A second problem with retrospective data is that possible causal relationships cannot be inferred.
13. In contrast to retrospective data, prospective data makes it possible to challenge theories or stories by comparing predictions to new data. Few demonstrations of the possibility of knowledge are more compelling than predicting otherwise improbable observations.
14. A distinction can be made between two types of prospective data: data that can be influenced by the researcher, and data that cannot be influenced. Influenced future data allows the manipulation of initial conditions, and so in principle allows the researcher to infer possible causality. If the researcher cannot manipulate experimental variables, then possible causal relationships cannot be inferred.
15. Whether one holds a theory-conserving or theory-discarding skeptical attitude should depend on the moral repercussions of making a false-positive or false-negative error. This risk will change from one claim/hypothesis/interpretation to the next.
16. Scholars in all fields of study ought to maintain flexibility in choosing a methodology that is suited to the task at hand. That choice should be informed by both the ethical repercussions of making various types of errors, as well as by the particular circumstances of the field itself.
17. In nearly every case, scholarship is enhanced by the availability of additional evidence. Like prosecuting attorneys, scholars have a moral obligation to seek out additional sources of evidence/data whenever these can be obtained. The magnitude of this obligation is proportional to the moral repercussions of the hypothesis.
18. Inferential statistical tests can be used equally effectively by both theory-conserving and theory-discarding skeptics. Theory-conserving skeptics have under-utilized statistical tests.
19. The material and structural conditions of any field of research are susceptible to change. A common source of change is either an increase or decrease in available pertinent data. Changing conditions often demand changes in research methodologies in order to minimize moral risks.
20. The selection of an appropriate methodology is a moral decision. When a scholar is unaware of the methodological choices, the selection of a methodology will be morally uninformed.
21. Research methodologies should be regarded as scholarly tools; researchers should resist the tendency to hold methodologies as comprehensive belief systems about the world.
22. There is no known methodological algorithm that ensures the advance of knowledge. Methodology consists primarily of a set of pointers that warn scholars of previously encountered pitfalls. Methodologies are extended and refined in the same manner as other theories.

[187] In this paper, I have endeavored to rekindle the view that the humanities are distinguished from the sciences primarily by their subject matter, and secondarily by a philosophical tendency towards humanistic rather than mechanistic conceptions of the world. More importantly, I have argued *against* the idea that the sciences and humanities are necessarily distinguished by their methodological habits. It is true that humanities disciplines currently tend to embrace false-negative skepticism, tend to be historical in orientation, tend to prefer pluralism to parsimony, and tend to prefer open accounts rather than closed explanations. However, I have noted that these methodological tendencies primarily arise from the structures and material circumstances attending the particular fields of study involved. Specifically, many humanities disciplines (though not all) are comparatively data-poor, deal with lower risk hypotheses, and are unable to carry out formal experiments. Data-poor disciplines repel false-positive skeptics because such disciplines provide an environment where false-positive skepticism is not productive.

[188] My claim that methodological differences arise primarily from the concrete research conditions of individual disciplines should evoke no surprise. Philosophers of knowledge all presume that what might loosely be called "rationality" is not discipline-specific. What is good for the epistemological goose ought to be good for the epistemological gander as well.

[189] Fields of study do have discipline-specific methodological needs. For example, manuscript studies have developed analytic methods based on water marks, chain lines, binding patterns, and so on. But there are also underlying patterns to how different disciplines approach their goals, and there are some unifying principles in research. In summary, while the humanities and sciences may rightly diverge in their philosophical conceptions about the nature of the world, they nevertheless share deep methodological commonalities. All fields of study can greatly benefit from an awareness of both the wide variety of available research methods and the innumerable pointers to methodological potholes.

The New Empiricism

[190] Research begins when we ask questions about the world. In the case of music, there is a multitude of worthwhile questions that can be posed. In many cases, there are negative moral repercussions if we choose *not* to investigate some question. Offering the excuse that "we could never be certain about the answer to that question" is hollow rather than noble, since it applies to all empirical questions. Good questions rightly challenge scholars to do our best to assemble evidence that might help produce informed (albeit limited and provisional) answers.

[191] Over the past decade, increasing numbers of music scholars have become attracted by the opportunities offered through empirical methods. The new empiricism recognizes that formal observation can indeed potentially lead to genuine insights about music and musicality. As I have noted, what the new empiricism shares in common with postmodernism is the conviction that scholarship occurs in a moral realm, and so methodology ought to be guided by moral considerations.

[192] Of course some research questions are hampered by a dearth of pertinent evidence. Nevertheless, there are reasonable ways of trying to decipher *likelihood* -- even if we can never divine the *Truth*. Many questions allow us to collect lots of pertinent data, and to use inferential statistical methods that allow us to minimize both false-positive and false-negative errors.

[193] The new empiricism has three bones to pick with the sciences. Scientists are wrong to denigrate or ignore fields that are data-poor and areas of research where experimentation is impossible. Scientists are wrong to treat the 0.05 confidence level as some sort of immutable inferential standard. For severely data-limited fields, 0.10 and 0.20 confidence levels ought to be entertained when the risks associated with making a false positive error are low. Scientists are also wrong to assume that the goal of research must always be to minimize false-positive errors.

[194] Similarly, the new empiricism also has some bones to pick with our colleagues in the humanities. Empiricism is not a dirty word. There are many musical questions, from history, esthetics, culture, analysis,

theory, performance, poeitics, reception, listening, etc. which can be usefully addressed using inferential statistical methods. Contrary to a popular belief, statistics cannot be used to prove any point of view.

[195] To the traditional music scholar, it must look for all the world like science is muscling-in on musicology. But the rise of empiricism has nothing to do with "science". It arises from within music scholarship, and is motivated by the desire to learn as much as possible from the information available to us -- including the additional information that might be assembled with a little effort. The pursuit of evidence is a moral obligation. Once again, the analogy to jurisprudence is compelling: if a prosecuting attorney has the opportunity to gain access to a wealth of new evidence, it would be morally reprehensible not to examine the material in order to better establish the guilt or innocence of someone.

[196] The pursuit of rigor is not some sort of methodological fetish. It is simply an attempt to avoid well-documented pitfalls in research. We ought not to be cynical of those scholars who aspire to do their best.

[197] In light of the above observations concerning methodology, it should be obvious that I think both humanities scholars and scientists should be educated with an aim to providing a broader repertoire of research methodologies. In particular, humanities scholars ought to learn the basics of statistical inference, and scientists ought to be exposed to phenomenological and deconstructionist approaches.

[198] Finally, moral and ethical philosophers should take a greater interest in epistemological ethics. Knowledge claims have consequences, and it is important for scholars to be cognizant of the moral and esthetic repercussions of their views -- including the view that something is unknowable. Better research on risk is needed in order to help researchers recognize when to adopt a theory-conserving or theory-discarding stance.

Footnotes

{1} It should be noted that the term "Positivism" is rarely used by modern empiricists; however, it is a designation commonly used in humanities scholarship, hence our use of it here. For a discussion of the so-called "culture wars" see: Alan Sokal and Jean Bricmont, *Fashionable Nonsense: Postmodern Intellectuals' Abuse of Science*, New York: Picador, 1998; and Joseph Natoli's *A Primer to Postmodernity,* Oxford: Blackwell Publishers, 1997 -- notably Chapter 8: Postmodernity's War with Science. Return.

{2} See Belsey (1993), Feyerabend (1975), Foucault (1970, 1977), Hartsock (1990), Kuhn (1962/1970), Natoli (1997). Return.

{3} In the pithy words of Foucault, "There is no power relation without the correlative constitution of a field of knowledge, nor any knowledge that does not presuppose and constitute at the same time power relations." (p. 27). Return.

{4} Throughout this article, the word "theory" should be interpreted broadly to mean any claim, hypothesis, theory, interpretation or view. Return.

{5} A standard textbook on scientific method notes the following: "In contrast to the consequences of publishing false results, the consequences of a Type II error are not seen as being very serious." (Cozby, 1989; p. 147). Return.

{6} It is essential to recommend *new* rather than established quacks. Established quackery has usually been the subject of research that has failed to establish its efficacy. Untested quackery has a better chance of being helpful. Return.

{7} Once again, the reader is reminded that throughout this article, the word "data" should be interpreted broadly to mean any information or evidence. Return.

{8} An example will be given later in this article. Return.

{9} "In musical interpretations, complexity is cherished ... In the social sciences, complexity seems to be avoided: the details of phenomena are levelled so that the findings can be expressed in the simplest possible way." (Rahn, 1983; p. 197).

{10} Statisticians have written extensively about the phenomenon of regression-toward-the-mean. Unfortunately, it appears to be a concept that is difficult for humans to grasp. Even Nobel laureate, W.F. Sharpe, incorrectly mistook regression-toward-the-mean for a new economic phenomenon (see, for example, Gary Smith, "Do Statistics Test Scores Regress Toward the Mean?"). As often happens with significant discoveries, a careful literature search sometimes finds that the same discovery was made decades earlier by another scholar. In a 1924 study, Henry Watt suggested that gap-fill in music can be attributed to regression toward the mean. Given the poor level of statistical numeracy among music scholars, I predict that it will take another 70 years before the preponderance of music theorists understand what has been demonstrated regarding gap fill. Return.

{11} There may be statistical reasons for excluding some data from an analysis. Return.

References

Agassi, Joseph. 1975.
Science in Flux. Dordrecht, Holland: D. Reidel Publishing Co.

Anderson, Philip. 1972.
More is different. *Science,* Aug. 4, 1972, pp.393-

Ayer, Alfred J. 1936
Language, Truth and Logic. London: Victor Gollantz.

Belsey, Catherine. 1993.
"Towards Cultural History," in: *A Postmodern Reader,* J. Natoli & L. Hutcheon (eds.) New York: SUNY Press, pp. 561, 555-556 [sic]. Quote is from page 561.

Cozby, Paul C. 1989.
Method in Behavioral Research. Mountain View, CA: Mayfield Publishing Co., 4th edition.

Cook, Nicholas. 1990.
Music, Imagination, and Culture. Oxford: Clarendon Press.

Duhem, Pierre. 1905.
La Théorie Physique: Son Objet, Sa Structure. Paris: Marcel Rivière; 2nd edition, 1914. Translated by P.P. Wiener as: *The Aim and Structure of Physical Theory* Princeton: Princeton University Press, 1954; reprinted New York: Atheneum, 1962, 1977.

Elster, Jon. 1989.
Nuts and Bolts for the Social Sciences. Cambridge: Cambridge University Press.

Feyerabend, Paul. 1975.
Against Method: Outline of an Anarchistic Theory of Knowledge. London: Verso Editions 7.

Fischer, David H. 1970.
Historians' Fallacies: Towards a Logic of Historical Thought. New York: Harper Collins.

Fitt, Mary. 1959.
Diels' Fragmente der Vorsokratiker. Translated by Kathleen Freeman as *Ancilla to The pre-Socratic philosophers: A complete translation.* Cambridge: Harvard University Press, 1983; 5th edition.

Foucault, Michel. 1970.
The Order of Things: An Archaeology of the Human Sciences. New York: Pantheon, 1970.

Foucault, Michel. 1977.
Discipline and Punish: The Birth of the Prison. New York: Vintage/Random House.

Gell-Mann, Murray. 1994.
The Quark and the Jaguar. New York: W.H. Freeman.

Gellner, Ernest. 1974.
Legitimation of Belief. Cambridge: Cambridge University Press.

Glashow, Sheldon & Ginsparg, Paul. 1986.
Desperately seeking superstrings. *Physics Today,* May 1986, p.7.

Gleick, James. 1987.
Chaos: Making a New Science. New York: Penguin Books.

Goldin, Claudia & Rouse, Cecilia. 2000.
Orchestrating impartiality: The impact of `blind' auditions on female musicians. *American Economic Review*, September 2000.

Goodman, Nelson. 1976/1981.
Languages of Art; An Approach to a Theory of Symbols, Brighton, Sussex: Harvester Press.

Gould, S.J. 1978.
Sociobiology: The art of storytelling. *New Scientist,* Vol. 80, pp. 530-533.

Gould, S.J., & Lewontin, R.C. 1979.
The spandrels of San Marco and the Panglossian program: A critique of the adaptationist programme. *Proceedings of the Royal Society of London*, Vol. 250, pp. 281-288.

Hacking, Ian. 1995.
Rewriting the Soul: Multiple Personality and the Sciences of Memory. Princeton, N.J.: Princeton University Press.

Hartsock, Nancy. 1990.
"Foucault on power: A theory for women?" In: Linda Nicholson (ed.) *Feminism/Postmodernism.* New York: Routledge.

Hattiangadi, Jagdish N. 1983.
A methodology without methodological rules. In: R. Cohen & M. Wartofsky (eds.), *Language, Logic, and Method.* Dordrecht, Holland: D. Reidel Publishing Co., pp. 103-151.

von Hippel, Paul. 2000a.
Redefining pitch proximity: Tessitura and mobility as constraints on melodic intervals. *Music Perception*, Vol. 17, No. 3, pp. 315-327.

von Hippel, Paul. 2000b.
Questioning a melodic archetype: Do listeners use gap-fill to classify melodies? *Music Perception*, Vol. 18, No. 2, pp. 139-153.

von Hippel, Paul, & Huron, David. 2000.

Why do skips precede reversals? The effect of tessitura on melodic structure. *Music Perception*, Vol 18, NO. 1, pp. 59-85.

Horgan, John. 1996.
The End of Science: Facing the limits of knowledge in the twilight of the scientific age. Reading, MA: Addison-Wesley Publisher.

Huron, David. 1996.
The melodic arch in Western folksongs. *Computing in Musicology*, Vol. 10, pp. 3-23.

Kuhn, Thomas S. 1962/1970.
The Structure of Scientific Revolutions. Chicago: University of Chicago Press.

Laudan, Larry. 1977.
Progress and Its Problems: Towards a Theory of Scientific Growth. Berkeley: University of California Press.

Lewontin, R.C. 1991.
Biology as Ideology. Concord, Ontario: Anansi Press and the Canadian Broadcasting Corporation.

Natoli, J. 1997.
A Primer to Postmodernity. Oxford: Blackwell Publishers.

Nettl, Bruno. 1985.
The Western Impact on World Music, New York: Schirmer Books.

Neyman, J. and Pearson, E.S. 1928
On the use and interpretation of certain test criteria for purposes of statistical inference. *Biometrika,* A, Vol. 20, pp. 175-240,263-294.

Neyman, J. and Pearson, E.S. 1967.
Joint Statistical Papers. Cambridge: Cambridge University Press.

Pagels, Heinz. 1988.
The Dreams of Reason. New York: Simon and Schuster.

Pascal, Blaise. 1669.
Pensées.

Plotkin, Henry. 1994.
Darwin Machines and the Nature of Knowledge. Cambridge, MA: Harvard University Press.

Poggioli, Renato. 1962/1968.
Teoria dell'arte d'avanguardia. Società editrice il Mulino. Translated by G. Fitzgerald as *The Theory of the Avant-Garde.* Cambridge, MA: Harvard University Press.

Polanyi, Michael. 1962.
Personal Knowledge: Towards a Post-Critical Philosophy. Chicago: University of Chicago Press.

Popper, Karl. 1934/1959.
Logik der Forschung. Vienna, 1934. Trans. as *The Logic of Scientific Discovery.* New York: Basic Books, 1959.

Rahn, John. 1983.
A Theory for All Music: Problems and solutions in the analysis of non-Western forms. Toronto: Toronto University Press.

Rahn, John. 1993.
"Differences" *Perspectives of New Music,* Vol. 31, No. 2, pp. 58-71.

Quine, Willard Van. 1953.
From a Logical Point of View. Cambridge, MA: Harvard University Press.

Roberts, Clayton. 1996.
The Logic of Historical Explanation. University Park, PA: Pennsylvania State University Press.

Rosen, D.E. 1982.
Telostean interrelationships, morphological function and evolutionary inference. *American Zoologist,* Vol. 22, pp. 261-273.

Seay, Albert. 1975.
Music in the Medieval World. Englewood Cliffs, NJ: Prentice-Hall Inc.

Watt, Henry J. 1924.
Functions of the size of interval in the songs of Schubert and of the Chippewa [i.e., Ojibway] and Teton Sioux [i.e., Lakota] Indians. *British Journal of Psychology,* Vol. 14, No. 4, pp. 370-386.

Webb, E.J., Campbell, D.T., Schwartz, R.D., Sechrest, R., and Grove, J.B. 1981.
Nonreactive Measures in the Social Sciences. Boston: Houghton Mifflin, 2nd edition.

 <!--As we learned in Lecture, psychology escaped the excesses of behaviorism with the advent of the cognitive revolution which re-introduced mental states as legitimate topics of investigation. It is perhaps ironic that computer technology played an important role in facilitating the acceptance of mental states. Although not easily observable, computer memories could clearly exist in a number of different states, and these states could significantly affect the ensuing computational behavior. /P>

University of California, Berkeley
Department of Music

The 1999 Ernest Bloch Lectures

Lecture 4. What is a Musical Feature?

Forte's Analysis of Brahms Opus 51, No. 1 Revisited

David Huron

Abstract

A theory of musical features is presented. The theory emphasizes how a given work is distinguished from other works in some musical corpus. The theory is illustrated by evaluating the principal motive proposed by Allen Forte in his analysis of Brahms's opus 51, no. 1. Forte's "alpha" interval-class set is shown to be unable to distinguish quartet No. 1 from other quartets by Brahms. On the basis of the theory, several refinements are made to Forte's alpha feature. Only when the prime form of the interval-class pattern is joined with a long-short-long rhythm does the resulting feature become distinctive of the work. Perceptually-pertinent properties are shown to assist in assembling a feature definition.

Imagine that you were robbed by someone and were later asked by police to provide a description of your assailant. Suppose you began your description by noting that the robber had a nose, two eyes, a mouth, and two ears. These "facts" would undoubtedly be "true," but the police would be rightly dismayed by your description for the simple reason that the facts fail to distinguish your assilant from a world of potential suspects.

When characterizing a musical work it is important, not simply to make accurate or truthful observations about its organization, it is also important to identify those characteristics that set the work apart from other musical works. In this paper, I propose to argue that music analyses can occasionally fall prey to such empty descriptive language. Like the description of our hypothetical robber, many otherwise truthful observations simply fail to be informative.

The paper is divided into two parts. Part One begins with a general theoretical discussion concerning the notion of a "feature." In the course of this discussion a number of properties of features will be distinguished and a set of corresponding terms defined. In Part Two, the concepts developed in Part One will be illustrated by referring to Allen Forte's analysis of the first movement of Brahms's string quartet, opus 51, no. 1. [1] In particular, the analysis will focus on what Forte identifies as the principal motivic feature of this work. Using the theoretical concepts related to features, several refinements will be made to Forte's motivic description. In effect, a `nose' will transformed into an `aquiline nose;' `two eyes' will be transformed into `close-set eyes.' A systematic analytic path will lead us to a motivic description that closely resembles a traditional diatonic motive as the principal intratextual feature of this work.

At the outset, it is important to reassure readers about the goals of this essay. My essay raises no criticisms whatsoever about set theory or the application of set theory to tonal music. [2] Indeed, set theory has made some

admirable strides in providing greater clarity in feature descriptions. Nor should the paper be regarded as an attack on Dr. Forte. [3] Forte's analysis simply provides a convenient illustration of a problem that is common in music analysis. Rather, my intention is to demonstrate that grievous problems of descriptive language are evident even in the work of exceptionally observant theorists -- and to show that such problems are (at least in principle) avoidable.

In keeping with a set-theoretical approach, my analytic points here will stress the musical foreground. Since set-based analytic methods have difficulty with background features (without recourse to Schenkerian or other models), my own analytic points will also leave background characterizations to other analytic methods. My focus on foreground features is merely a matter of convience and of avoiding undue length. As the reader will readily understand, the inferential analytical approach employed here is as pertinent to the comparative evaluation of proposed background, middleground, and process-related features as it is to foreground motivic features.

PART ONE: WHAT IS A FEATURE?

In common parlance, a "feature" is a notable or characteristic part of something; a feature is something that helps to distinguish one thing from another (or one group of things from another group of things). In the case of the analysis of artistic artifacts, features provide essential descriptive primitives by which the unique characters of individual works may be identified, and by which the commonalities of styles, periods, and genres may be portrayed.

Properties of Features

Broadly-speaking, two closely-linked classes of features can be distinguished: *intratextual* and *intertextual.* Intertextual features arise from relationships between works, whereas intratextual features arise within the context of the work itself. At a minimum, we might expect a proposed feature to be *present* in an artifact or class of artifacts. If it is claimed that *X* is a feature of work *Y,* then we would naturally expect *X* to be located, or evident in *Y*. Sometimes features may be notable by their absence, however. For example, the absence of harmonic thirds and sixths in a tonal-period work would be something of an aberration. Since, the absence of something commonplace may itself be noteworthy, we might refer to this property as *negative presence.* Note that negative presence presupposes the existence of a normative repertoire or established set of expectations. That is, one can't recognize that something is missing unless one has a sense of what is normally present.

As in the case of positive presences, negative presences may be either intertextual or intratextual. An intertextual negative presence arises when the accumulated evidence suggests that a work belongs to a certain class of works, yet fails to exhibit a property that is otherwise assumed to be essential to membership in that class or group. By contrast, an intratextual negative presence arises within a work when the work itself establishes the expectation of some event or phenomenon that nevertheless remains unrealized. Since the thwarting of an expectation may quickly become a cliché, intratextual features are constantly being transformed into intertextual norms. [4]

Note that even in the case of positive presences, a proposed feature is rarely directly represented in an artistic artifact. For example, the pitches of a score do not directly encode the melodic intervals. The concept of "melodic interval" relies on the assumption of an underlying "voice" or "part" and deciphering voicing sometimes entails remarkably sophisticated interpretations. On what basis, then, can one defend the assumption of *voice?* Those theorists who have contemplated such matters typically rely on one of two appeals. One might appeal to notational conventions such as the use of separate staves or differentiation via stem direction. A more common appeal is to the perceptual experiences that affirm the subjective phenomenon of `musical line' and hence of `melodic interval.'[5] We often assume that the notational conventions of stem directions and independent staves are straightforward reflections of a common psychological experience. That is, theorists rightly consider melodic intervals to be implied by individual successive pitches, and rightly assume that such intervals may be readily derived in most circumstances. [5]

In fact, few features can be identified in the "raw data" of an artifact. Like melodic intervals, most features are deciphered, interpreted or derived from the raw data. Of course even the "raw data" are interpreted. What makes the data seem ontologically "raw" is that the interpretations are comparatively stable and uncontentious. We can continue to use such terms as "raw data" as a linguistic convenience, although the term should be understood as a short-hand for interpretations that are less contentious -- and the term should never be seen as a foundation that is immune to further conceptual analysis or challenge.

Of course derived presences may be much more indirect than is the case for melodic intervals. Derived presences may include such concepts as deceptive cadences, prolongations, syncopations, suspensions, ritornellos, and other phenomena. Often, these derived concepts are more notable and compelling features than directly notated pitches or durations.

Note that there are an infinite number of possible derived presences, and not all such derivatives can have the same analytic status. For example, one might rigorously define the property *fuddle* to be the semitone pitch distance between the second-last notes occurring in odd numbered measures whose stem directions extend upward. Implicit in any theory is the idea that certain properties (such as melodic intervals) are not reified derivative concepts, and that, unlike the concept of *fuddle,* they represent some organizing principle that has influenced either the construction of the work, or its reception, or both. In the case of conventional tonal theory, the notion of a diatonic step is an example of such a central concept. In the case of set theory, the notion of an interval-class is an example of such a derivative concept.

In discussing above how theorists might justify the traditional *voice*-assumption underlying the concept of the melodic interval, two different forms of support were cited. One was based on subjective perceptual experience, and the other was based on notational practice. It is important to understand that these are just two of innumerable forms of appeals. One need not refer, either to perception or to the notation, in order to justify a useful descriptive concept. Consider, by way of example, the notion of something being "idiomatic." In works written for trumpet, it has been observed that difficult finger/valve combinations are systematically less common in works written by trumpet virtuosi than in trumpet works written by non-virtuosi. [6] In most circumstances, it is only the experienced trumpet player who is able to apprehend the specific properties that distinguish an idiomatic arrangement from an unidiomatic arrangement. That is, a trumpet player may experience a passage according to such categories as "easy to finger" or "difficult to finger." Moreoever, these conceptual and descriptive categories can be defended, even if listeners are unable to perceive them, even if non-trumpet performers fail to experience them, and even if theorists are oblivious to their existence.

Descriptive languages are often unique to particular communities. Usually, the language is largely given *a priori,* but there are always opportunities to expand the descriptive conceptual vocabulary. A composer might choose to create a work whose organization is intimately linked to the *fuddle* concept. If this were the case, the organization would have salience only for readers of this article -- and even then, only if the reader has access to the notated score. This simply illustrates that useful descriptive concepts can spring into existence within a given community, or culture.

When we describe features, we are at liberty to choose the descriptive language, although the language must contain concepts whose assumptions have some degree of support. In describing a robber to the police, an optician might focus on the robber's brand of eye-glasses, a fashion designer might focus on features of the robber's clothing, and a dialect expert might focus on the robber's manner of speech. However, within each of these descriptive domains, it is still possible to give either fruitful or empty descriptions. The optician's recognition a common brand of sun-glasses, the fashion designer's notice of poor color-coordination, and the dialect expert's animated description of the robber's impolite vocabulary may be of little use in distinguishing our robber from from anyone else.

The mere presence of some element or property does not necessarily make it a good feature. A good feature must in some ways draw attention to itself. It must be notable, or what might be dubbed *salient.* Of course there are innumerable ways by which something may draw attention -- some of which we will discuss later. One of the most venerable properties is the simple *prevalence* of an event; that is, how frequently a pattern recurs.

Apart from such intratextual factors, there are also intertextual factors that can contribute to salience. A single statement of "B-A-C-H" or the opening notes of *Dies Irae* can be sufficient to establish the notability of the event. Intertextual factors may be either intentional or unintentional on the part of the artist. Intentional factors may include *quotation, allusion, parody* and *model* -- concepts that have received some lucid theoretical attention. [7] To these may be added the concept of *evocation* where a passage unintentionally reminds the listener of a similar passage in another work. (Since evocation is defined as unintentional, the clearest examples occur where a musical passage reminds us of a similar passage in a work created much later by another composer.)

To summarize, we may define salience as a heightened attention that can arise due to either intratextual factors (such as phenomenal accent) or intertextual factors (such as quotation).

In identifying a feature, we must always be cognizant of the question "feature of what?" Features may be characteristic of a work, of a movement, of a composer, of a style, of tonal music in general, and so forth. What constitutes a feature depends on the scope of our gaze. For example, a non-Western listener may deem the sound of the pianoforte to be a significant feature of Western music -- without distinguishing any further divisions. To the music analyst, the features of principal interest have been those that characterize individual works, or sections thereof. At the supra-opus level, features of interest may include stylistic features, and the three basic lines of Schenkerian analysis, etc.

Evaluation of Features

All scholarly disciplines are founded on some type of descriptive enterprise. Features provide the essential building-blocks for such descriptions. Each discipline establishes its own criteria for good description, however, some evaluative criteria are nearly universal and possibly tautological. Perhaps the most important property of a description is the degree to which it distinguishes the objects of our attention from other (possibly like) objects. If the goal of an analytic description is to convey what is unique or distinctive of a given object or class of objects, then good features must embody or define some of that distinctiveness.

Another important property of feature descriptions is the economy of the means of expression. When asked for directions by a passing motorist, the success of our description depends not only on the distinctiveness of the landmarks we mention, but also -- since human memory is fallible -- on the brevity of our description. Of course, in giving directions to our motorist, we might ask if they are already familiar with certain landmarks. In inquiring about their existing knowledge, we are, in effect, determining whether intertextual references will have any value in enhancing the description. Whether or not our motorist has knowledge of the pertinent geography, a good description will identify a succinct set of distinctive features. [8]

In the absence of distinctiveness, many otherwise truthful characterizations are without merit due to the excessive breadth of the description. In concrete terms, we may define *distinctiveness* as the property of "relative salience" -- that is, where a feature is more characteristic of the artifact being described compared with other artifacts or phenomena. Note that the property of distinctiveness is necessarily comparative.

Once again, what constitutes a distinctive feature depends on the scope of our gaze. A "nose" is a feature of faces in general, but the presence of a nose cannot be a feature of a particular face. It is the `stubby nose' or the `pixie nose' that may be a feature of some given face -- but only because of the implied comparison with other noses.

All of the foregoing attributes or properties of a feature are attempts to approach the central issue related to a feature: its importance, eminence or significance. By their very nature, the concepts of importance and significance are open-ended. A single, apparently trivial detail may be the source of a striking story. Significance is always contextual, and there is no telling the historical, social, personal, formal or other domains that establish the context for some event -- however apparently minor or trivial. This means that it is impossible, in principle, to produce an exhaustive inventory of the significant features of some work.

Although we cannot describe *all* the potentially significant features of a work, or even *the* most significant features of a work, this does not preclude the possibility of identifying *some* features that are significant. My claim here is that, salient distinctive features will always bear some degree of significance in a work -- even though other features may exist that can claim to be of greater significance.

By way of summary, the feature-related properties we have discussed are reviewed in Table 1. In general, a good feature may be defined as a succinctly described characteristic that is salient, distinctive, and significant. That is, a good feature is something that attracts our attention in some way, that distinguishes an object or class of objects from other like objects, and that is worthy of interest.

Table 1

Summary of Terminology Related to Features

General Terms:	
presence	existent within an artifact
negative presence	absent from an artifact (although expected)
salience	noticeability of an event
distinctiveness	greater salience compared to occurrences in other artifacts
significance	notable, important, worthy of attention
Some Intratextual factors contributing to salience:	
prevalence	noticeable because it recurs frequently
accent	noticeable because it is stressed (e.g., dynamic, agogic ...)
recency	noticeable because it is last
primacy	noticeable because it is first
mnemonic	noticeable because it is easily remembered
Some Intertextual factors contributing to salience:	
evocation	unintended reminder of similar passage in another artifact
quotation	intended exact quotation from another artifact
allusion	intended indirect reference to another artifact
parody	intended exact or indirect reference, intended to spurn
model	intended or unintended borrowing of a structural framework

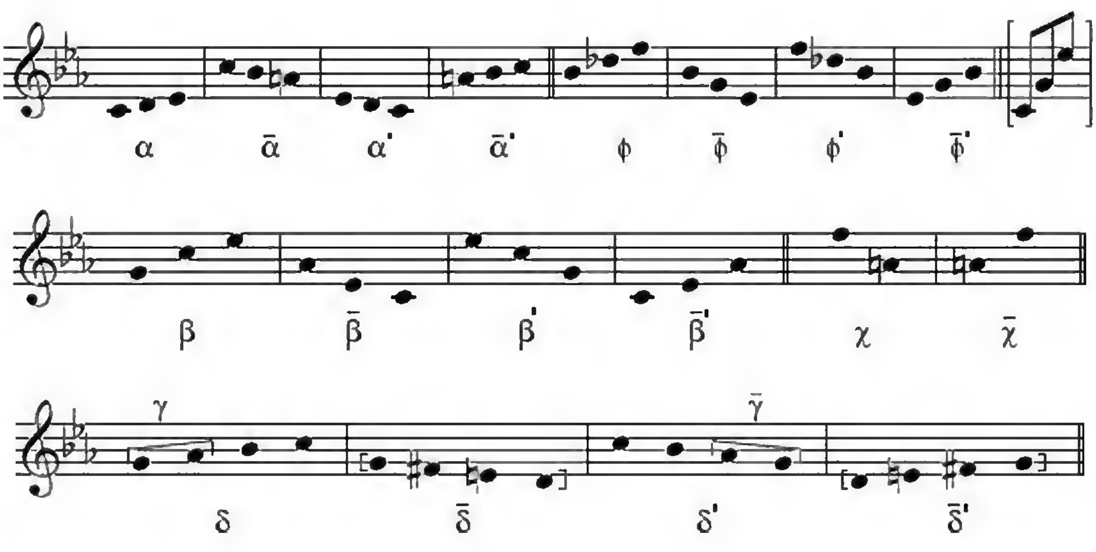
PART TWO: FEATURES IN BRAHMS'S OPUS 51, NO. 1

The conceptual framework outlined in Part One may be better understood through an illustrative example. Example 1 shows the opening measures of the first movement of Brahms's opus 51, no. 1. The opening statement in the first violin part constitutes what would normally, and informally, be dubbed the principal theme of this movement -- a theme that Hanslick characterized as `magnificently passionate.' [9] Example 2 shows a number of interval-related groups that are identified in Forte's analysis as playing a central role in this movement. The most important of these sets is the interval sequence (+2,+1) which, along with its inversion (-2,-1), retrograde (+1,+2), and retrograde inversion (-1,-2), Forte dubs the `alpha' group of motives. In order to distinguish the interval patterns from the interval-class patterns we will use the unsigned designation (2,1) to refer to the interval-class that embodies all four of these individual interval patterns.

Example 1.

Brahms, string quartet Op. 51, No. 1, mov.1, mm. 1-6.

Example 2.



Some sample sets from Forte's analysis of Brahms Op. 51, No. 1, mov. 1.

There is no question that the alpha pattern is *present* in opus 51, no. 1. The pattern occurs in the opening three notes of the first violin and appears numerous times throughout the movement. According to our analytic taxonomy, we need to consider the degree to which occurrences of this pattern are *salient* and *distinctive.*

Of the various factors contributing to salience, consider first the simple property of *prevalence.* Out of 7,045 interval-class melodic diads found in this movement, the alpha pattern is the most common interval-class pattern, occurring 352 times. (The prime form alone occurs 136 times making it the 13th most common two-interval pattern.) [10]

However, since scale-like movements of tones and semitones are ubiquitous in tonal music, we might well expect the alpha pattern to be common in most musical works. Returning to the police station, suppose that we remain unconvinced by the police officer's protestations that saying our assailant "has a nose" is utterly vacuous. The officer might pull out a college yearbook and ask us to count the number of photographs where a person is without a nose. Having counted a large number of people, each with a nose, we might be more convinced that our description does indeed fail to be distinctive. Such a statistical demonstration might seem crude or unwarranted, but it serves the important purpose of illustrating the shortcomings of our description.

In order to determine whether the alpha pattern is *distinctive* of the first movement of opus 51, no. 1, we need to compare this movement with other musical works. But what other works would provide an appropriate comparison? If we compared Brahms's opus 51 to (say) a Brazilian samba, [11] then the origins of any observed differences between the two pieces would be difficult to interpret. Perhaps the differences would arise due to different historical periods, or different nationalities of the composers, or different instrumentation, or different styles. In other words, the observed differences may be manifestations of different classes of compositions (such as genres) rather than differences in the works themselves.

In order to minimize these interpretive problems it is preferable to select comparison works that are as similar as possible to the work we are attempting to describe. A good sample of music for such a comparison is all of Brahms's remaining string quartet movements. Brahms claimed to have written over 20 string quartets in his youth -- works which he later destroyed. Of his mature works, Brahms published three string quartets: opus 51, no. 1; opus 51, no. 2; and opus 67. These works were written over a period of about a decade. By choosing this repertoire, we can be confident in dismissing claims that any differences we may observe are due to different composers, nationalities, genres, styles, instrumentation, etc. That is, any observable differences are more likely attributable to the different characters of the *works.* Of course the string quartet movements are not exactly "matched" -- there remain differences. The various movements exhibit different tempi, moods, and forms. Rather than comparing the first movement of Quartet No. 1 with the other movements of the same work, in some ways, a better comparison might be to compare the first movements from all three quartets (No. 1: Allegro (3/2 meter), No. 2: Allegro no troppo (2/2 meter), and No. 3: Vivace (6/8 meter)). The different meters and the different keys mean that we cannot dismiss the possibility that any observed differences can be attributed to these factors.

Table 2 tabulates the number of occurrences of each of the alpha patterns for each of the first movements of Brahms's three string quartets. Separate results are shown for all four set variants of Forte's interval-class motive.

Table 2

Prevalence of `Alpha' Patterns in First Movements of Brahms's String Quartets

interval -------------		Brahms Quartets ------------	
pattern	No. 1	No. 2	No. 3
+2,+1	136	72	139
-2,-1	94	129	226
-1,-2	52	116	199
+1,+2	70	104	110
	——	——	——

	352 (5.00%)	421 (6.37%)	594 (7.92%)
of	(7045)	(6612)	(7498) two-interval instances

The results of Table 2 show that the alpha interval-class pattern is prevalent in all three quartet movements -- not just in the first movement of quartet No. 1. In fact, the alpha interval-class pattern is proportionally more common in each of the other string quartet movements than it is in quartet No. 1. This implies (but does not prove) that the alpha interval-class is a musical commonplace -- at least in the case of Brahms, but probably for tonal music in general. In summary, we can say that the alpha interval-class pattern is *present* and *prevalent* in quartet No. 1, but, on the basis of this single salience factor, we cannot say that alpha is *distinctive* of quartet No. 1.

As we noted earlier, a good feature must draw attention to itself in some way. Apart from simple repetition (prevalence), there are innumerable other ways to draw attention. If we trust Forte's intuition, we must conclude that more than mere prevalence contributed to Forte's selection of the alpha motive. We need to look for other possible salience-enhancing properties. One place to begin is by noting that Forte selected the initial three notes of the work. In memory and learning tasks, psychologists have established that the first visual or auditory items in a sequence are more salient than other items -- a phenomenon dubbed *primacy.* (Similarly, people are better at recalling and learning items toward the end of a sequence -- a phenomenon known as *recency*.)

We might also note that Forte chose the upper-most voice, rather than, say, the initial notes in the viola part. We can only guess as to the source of Forte's intuition, but it is suggestive that this choice also accords with experimental research showing that listeners find outer voices, especially the upper-most voice, more perceptual salient. While we are considering these salience-enhancing treatments, lets consider some other common ones. On the most superficial level, salience can be enhanced by increasing the amount of acoustical energy (e.g., loudness and/or duration) that attends some event. All types of foreground accent (dynamic, agogic, melodic, etc.) can contribute to salience.

Perhaps it is the case that alpha tends to occur in foreground or melody parts rather than in accompaniment contexts. Motivic statements might tend to be presented via unison or octave doublings, or appear in outer-voice parts. Various forms of accent may be applied; statements may coincide with metrically strong positions. The feature might be isolated from preceding and ensuing material; it might tend to appear at the beginning and ending of the work, or at the beginnings and ends of phrases.

Tables 3a-f show the results of six exploratory analyses that attempt to determine the extent to which Forte's alpha patterns are linked with various contextual treatments that may be expected to enhance salience. Table 3a shows the number of instances of the alpha patterns that involve pitch-class doublings (such as unison or octave statements). Table 3b shows the number of instances of the alpha patterns that follow a rest (implying perceptual primacy). Table 3c shows the number of instances of the alpha patterns that precede a rest (implying perceptual recency). Table 3d shows the number of instances of the alpha patterns that coincide with the beginning of a phrase or slur mark (also implying primacy). Table 3e shows the number of instances of the alpha patterns that begin on the strongest metric position (down-beat) in a measure. Table 3f shows the number of instances of the alpha patterns that occur in outer voices ('cello and first violin). In each case, the results for opus 51, no. 1 are contrasted with similar tallies for the first movements of Brahms's other two string quartets.

Table 3a

Instances of `Alpha' Patterns Involving Pitch-class Doubling

interval	-------------	Brahms Quartets	------------
pattern	No. 1	No. 2	No. 3
+2,+1	11 (34th)	6 (39th)	12 (30th)
-2,-1	0	1 (95th)	32 (7th)
-1,-2	2 (85th)	3 (64th)	31 (10th)

+1,+2	9 (42nd)	5 (51st)	11 (32nd)
	____	____	____
	22 (1.72%)	15 (1.83%)	86 (7.26%)
of	(1282)	(820)	(1184) pitch-class-doubled instances
of	(120)	(112)	(108) unique pitch-class-doubled interval patterns

Table 3b

Instances of `Alpha' Patterns Following a Rest

interval -------------		Brahms Quartets ------------	
pattern	No. 1	No. 2	No. 3
+2,+1	24 (2nd)	9 (9th)	17 (3rd)
-2,-1	13 (9th)	11 (8th)	7 (10th)
-1,-2	4 (20th)	15 (3rd)	11 (4th)
+1,+2	1 (58th)	19 (1st)	10 (6th)
	____	____	____
	42 (11.83%)	54 (15.84%)	45 (17.24%)
of	(355)	(341)	(261) rest-linked instances
of	(75)	(88)	(71) unique rest-linked interval patterns

Table 3c

Instances of `Alpha' Patterns Preceding a Rest

interval -------------		Brahms Quartets ------------	
pattern	No. 1	No. 2	No. 3
+2,+1	18 (3rd)	14 (4th)	11 (1st)
-2,-1	4 (24th)	18 (2nd)	9 (3rd)
-1,-2	2 (36th)	14 (4th)	5 (10th)
+1,+2	3 (29th)	10 (7th)	1 (55th)
	____	____	____
	27 (7.56%)	56 (16.23%)	26 (9.89%)
of	(357)	(345)	(263) rest-linked instances
of	(78)	(102)	(85) unique rest-linked interval patterns

Table 3d

Instances of `Alpha' Patterns Coinciding with Slur or Phrase Onsets

interval -------------		Brahms Quartets ------------	
pattern	No. 1	No. 2	No. 3
+2,+1	27 (2nd)	35 (2nd)	22 (13th)
-2,-1	30 (1st)	12 (16th)	2 (79th)
-1,-2	4 (42nd)	4 (63rd)	36 (6th)
+1,+2	13 (14th)	41 (1st)	10 (26th)
	____	____	____
	74 (11.24%)	92 (9.27%)	70 (6.66%)

of	(658)	(992)	(1053) slur/phrase-linked instances
of	(118)	(220)	(138) unique slur/phrase-linked interval patterns

Table 3e

Instances of `Alpha' Patterns Beginning a Measure

interval pattern	Brahms Quartets No. 1	No. 2	No. 3
+2,+1	42 (2nd)	18 (9th)	7 (38th)
-2,-1	17 (9th)	12 (20th)	29 (9th)
-1,-2	6 (24th)	9 (28th)	23 (12th)
+1,+2	12 (13th)	6 (39th)	17 (19th)
	77 (9.55%)	45 (4.53%)	76 (7.20%)
of	(806)	(994)	(1056) downbeat-linked instances
of	(118)	(165)	(158) unique downbeat-linked interval patterns

Table 3f

Instances of `Alpha' Patterns In Outer-Most Voices

interval pattern	Brahms Quartets No. 1	No. 2	No. 3
+2,+1	79 (5th)	35 (12th)	71 (12th)
-2,-1	60 (7th)	50 (6th)	110 (6th)
-1,-2	33 (20th)	49 (7th)	86 (11th)
+1,+2	50 (9th)	42 (8th)	63 (13th)
	222 (6.68%)	176 (5.50%)	330 (8.53%)
of	(3321)	(3200)	(3868) outer-most pitch interval-diad instances
of	(341)	(385)	(327) unique outer-most pitch interval-diad patterns

Each table entry identifies the number of instances found, followed (in parentheses) by the rank position of that pattern in a list of all patterns in the movement -- ordered by frequency of occurrence. Summary statistics for each of the three quartet movements are provided at the bottom of each table. Specifically, the total number of alpha-related patterns are tabulated, followed (in parentheses) by the percent occurrence. The total numbers of pertinent measures are also identified followed by measures of the total number of unique patterns conforming to the defined context.

Consider first the summary statistics in each table showing the totals for all four interval patterns. In particular, compare the percentage occurrences for all three quartet movements. Tables 3a, 3b, 3c and 3f do not show any comparative coincidence between occurrences of alpha and pitch-class doubling, adjacent rest boundaries, or outer-voice positions. For example, Table 3a shows that only 1.72% of the alpha interval-class occurrences involved pitch-class doublings, whereas quartets No. 2 and No. 3 exhibited 1.83% and 7.26% respectively. However, Tables 3d and 3e appear to show a heightened occurrence of the alpha interval-class motive coinciding with slur or phrase onsets and beginning in down-beat metric positions -- compared to the other two quartets.

The results shown in Tables 3a-f are more telling if we focus separately on the four individual interval patterns subsumed by Forte's alpha interval-class pattern, Tables 4a-d recast the data in Tables 3a-f comparing the results for each of the string quartet movements for each of the four interval patterns: (a) the prime interval-pattern (+2,+1), (b) the inversion (-2,-1), (c) the retrograde (-1,-2), and (d) the retrograde inversion (+1,+2). Only the prime form (+2,+1) in Table 4a shows a significant increased presence compared with the comparison movements. [12] With the exception of the pitch-class doubling context, the (+2,+1) interval-pattern is linked with each of the examined salience-enhancing contexts much more often than for the other quartets. That is, the (+2,+1) interval pattern is more likely to be preceded or followed by a rest, is more apt to coincide with the beginning of a slur or phrase mark, is more apt to begin a measure, and is more apt to appear in an outer voice.

Table 4a

Comparison of the Salience of the (+2,+1) Interval Patterns

	----------------	Brahms Quartets	---------------
	No. 1	No. 2	No. 3
Table 3a	11/1282 (0.9%)	6/820 (0.7%)	12/1184 (1.0%)
Table 3b	24/355 (6.8%)	9/341 (2.6%)	7/261 (2.7%)
Table 3c	18/357 (5.0%)	14/345 (4.1%)	11/263 (4.2%)
Table 3d	27/658 (4.1%)	35/992 (3.5%)	22/1053 (2.1%)
Table 3e	42/806 (5.2%)	18/994 (1.8%)	7/1056 (0.7%)
Table 3f	79/3321 (2.4%)	35/3200 (1.1%)	71/3868 (1.8%)
	-----------	-----------	-----------
	201/6779 (3.0%)	117/6753 (1.7%)	130/7422 (1.8%)

Table 4b

Comparison of the Salience of the (-2,-1) Interval Patterns

	----------------	Brahms Quartets	---------------
	No. 1	No. 2	No. 3
Table 3a	0/1282 (0.0%)	1/820 (0.1%)	32/1184 (2.7%)
Table 3b	13/355 (3.7%)	11/341 (3.2%)	7/261 (2.7%)
Table 3c	4/357 (1.1%)	18/345 (5.2%)	9/263 (3.4%)
Table 3d	30/658 (4.6%)	12/1053 (1.1%)	2/1053 (0.2%)
Table 3e	17/806 (2.1%)	12/994 (1.2%)	29/1056 (2.7%)
Table 3f	60/3321 (1.8%)	50/3200 (1.6%)	110/3868 (2.8%)
	-----------	-----------	-----------
	124/6779 (1.8%)	104/6753 (1.5%)	189/7422 (2.5%)

Table 4c

Comparison of the Salience of the (-1,-2) Interval Patterns

	----------------	Brahms Quartets	---------------
	No. 1	No. 2	No. 3
Table 3a	2/1282 (0.2%)	3/820 (0.4%)	31/1184 (2.6%)
Table 3b	4/355 (1.1%)	15/341 (4.4%)	11/261 (4.2%)
Table 3c	2/357 (0.6%)	14/345 (4.1%)	5/263 (1.9%)

Table 3d	4/658 (0.6%)	4/1053 (0.4%)	36/1053 (3.4%)
Table 3e	6/806 (0.7%)	9/994 (0.9%)	23/1056 (2.2%)
Table 3f	33/3321 (1.0%)	49/3200 (1.5%)	86/3868 (2.2%)
	-----------	-----------	-----------
	51/6779 (0.7%)	94/6753 (1.4%)	192/7422 (2.6%)

Table 4d

Comparison of the Salience of the (+1,+2) Interval Patterns

	----------------	Brahms Quartets	---------------
	No. 1	No. 2	No. 3
Table 3a	9/1282 (0.7%)	5/820 (0.6%)	11/1184 (9.3%)
Table 3b	1/355 (0.3%)	19/341 (5.6%)	10/261 (3.8%)
Table 3c	3/357 (0.8%)	10/345 (2.9%)	1/263 (0.4%)
Table 3d	13/658 (2.0%)	41/1053 (3.9%)	10/1053 (0.9%)
Table 3e	12/806 (1.5%)	6/994 (0.6%)	17/1056 (1.6%)
Table 3f	50/3321 (1.5%)	42/3200 (1.3%)	63/3868 (1.6%)
	-----------	-----------	-----------
	88/6779 (1.3%)	123/6753 (1.8%)	112/7422 (1.5%)

Tables 4a-d suggest that it is not the alpha interval-class pattern that is distinctive of the first movement of Brahms's opus 51, no. 1; rather the results suggest that it is the interval-specific form (+2,+1) that is distinctive. This is not to suggest that motivic inversions (for example) don't appear in the movement. (Salient inverted statements clearly appear in measures 92-105.) It is only to say that the inversion is a rare variation of a more basic feature. It is appropriate then that Forte identifies the interval-specific pattern (+2,+1) as the *prime* form of the interval-class set -- and uses this term in preference to the proper normal form (1,2). However, the retrograde, inversion, and retrograde inversion forms lack distinctiveness in this movement.

If we wish to improve upon Forte's alpha feature, we need to consider how the feature description can be modified or extended such that it will truly become clearly *distinctive* rather than merely *prevalent.* There are a number of candidate ideas for modifying the feature description. One of the most important candidates is to consider the relationship between pitch and duration. For example, the three notes of Forte's alpha motive seem to be linked to a long-short-long rhythm. This suggests that we re-analyse the work with regard to patterns consisting of both interval diad and relative duration. We might begin by first analyzing the long/short duration patterns in the three first movements in order to see if the long-short-long pattern is present, prevalent, and distinctive of opus 51, no. 1. All of the rhythmic-diad patterns are shown in Table 5.

Table 5

Second Order Delta-duration Patterns in Brahms String Quartets

Duration Pattern	-------------	Brahms Quartets	------------
	No. 1	No. 2	No. 3
same same	3762	2533	3702
(65.29%)	(51.23%)	(62.41%)	
longer longer	18	80	34
shorter shorter	50	72	65
	-----	-----	-----
	68 (1.18%)	152 (3.07%)	99 (1.67%)

shorter longer	600	402	295
longer shorter	722	688	630
	-----	-----	-----
	1322 (22.94%)	1090 (22.05%)	925 (15.59%)
same longer	257	498	508
longer same	96	145	122
shorter same	205	394	452
same shorter	52	132	124
	-----	-----	-----
	610 (10.59%)	1169 (23.64%)	1206 (20.33%)
	-----	-----	-----
	5762	4944	5932

In Table 5, the absolute duration of the notes has been of no concern; rather, a duration is deemed to be long or short depending upon the length of the preceding note. Thus the patterns half-quarter-whole and half-eighth-quarter would both be deemed long-short-long rhythms despite the fact that the quarter-duration is deemed `short' in the first pattern and `long' in the second pattern. Of course, this is only one of many possible approaches to characterizing rhythmic patterns.

Table 5 indeed appears to suggest that the long-short-long pattern (i.e. "shorter-longer") is more prevalent in the first movement of string quartet No. 1 than in the first movements of Brahms's two other quartets. The pattern appears 600 times out of 5,762 contiguous 3-note instances in quartet No. 1 (i.e. 10.4% of all instances). The corresponding percentages for quartets Nos. 2 and 3 are 8.1% and 5.0% respectively. We now need to consider possible relationships between this potential rhythmic feature and the alpha interval pattern.

Tables 6a-c show the durational contexts of the alpha interval patterns for each of the first movements of the three Brahms string quartets. Some patterns (such as longer-longer and shorter-shorter) are notably rare. The important question is whether the alpha patterns in quartet No. 1 are more strongly linked to the shorter-longer (i.e. long-short-long) rhythmic contour. The answer is a resounding yes. But once again, the answer must be qualified to say that only the interval-specific pattern (+2,+1) is strongly linked to the shorter-longer durational pattern. The remaining interval-specific patterns display no more rhythmic linkage than is found in the other quartets.

Table 6a

Durational Context of `Alpha' Patterns for Brahms Quartet No. 1 (1st movement)

duration patterns	-------	Interval	Patterns	-------
	(2,1)	(-2,-1)	(-1,-2)	(1,2)
same same	22	41	18	31
same longer	16	0	6	18
same shorter	2	0	5	0
longer same	0	2	6	0
shorter same	2	0	0	0
longer longer	0	0	0	0
shorter shorter	0	0	0	2
shorter longer	44 (47.8%)	7 (10.9%)	4 (9.8%)	3 (5.2%)
longer shorter	6	14	2	4
	____	____	____	____

	92	64	41	58

Table 6b

Durational Context of `Alpha' Patterns for Brahms Quartet No. 2 (1st movement)

duration patterns -------		Interval	Patterns	-------
	(2,1)	(-2,-1)	(-1,-2)	(1,2)
same same	9	24	9	36
same longer	20	26	25	2
same shorter	0	4	0	0
longer same	0	4	9	0
shorter same	1	0	5	3
longer longer	0	6	7	0
shorter shorter	2	0	0	0
shorter longer	7 (14.0%)	25 (26.6%)	6 (8.1%)	3 (5.3%)
longer shorter	11	5	13	13
	____	____	____	____
	50	94	74	57

Table 6c

Durational Context of `Alpha' Patterns for Brahms Quartet No. 3 (1st movement)

duration patterns -------		Interval	Patterns	-------
	(2,1)	(-2,-1)	(-1,-2)	(1,2)
same same	45	125	116	51
same longer	5	12	3	1
same shorter	2	7	4	8
longer same	3	6	0	1
shorter same	2	3	0	2
longer longer	0	0	1	0
shorter shorter	9	1	1	0
shorter longer	25 (22.9%)	34 (16.9%)	25 (14.5%)	31 (31.6%)
longer shorter	18	13	23	4
	____	____	____	____
	109	201	173	98

An association or link between two parameters may be either unidirecitonal or bi-directional. For example, if we find that people with *X*-colored hair tend to have *Y*-colored eyes, it does not necessarily follow that people with *Y*-colored eyes tend to have *X*-colored hair. When two characteristics are reciprocally linked they form a bi-direction association. Bi-directional associations are especially noteworthy since they imply a common source or origin for both characteristics or parameters. The preceding tables (6a-c) established only that the (+2,+1) interval pattern has a strong tendency to occur in a long-short-long (shorter-longer) durational context. We should also determine whether shorter-longer durational patterns tend to coincide with the (+2,+1) interval context.

Table 7 shows the dozen most prevalent interval patterns exhibiting the shorter-longer durations in the first movement of quartet No. 1. As can be seen, the interval-specific (+2,+1) pattern is ranked first. The other set

variants for alpha are not shown in Table 7 due to their especially low rankings -- 27th (-2,-1), 42nd (-1,-2), and 58th (+1,+2). In summary, Tables 6a and 7 show that the bond between the interval pattern (+2,+1) and the long-short-long rhythm is both strong and bi-directional. This linkage suggests that the interval and rhythm attributes share a common origin and are inextricable components of a *single* pattern. This finding suggests that it would be inappropriate to define the interval feature independent of its rhythmic properties. By contrast, none of the other set variants show any such rhythmic linkage. For example, a separate analysis of the retrograde forms of alpha showed no correlation with retrograde forms of the rhythm.

Table 7

Interval Dyad Contexts for `Shorter-Longer' Durational Patterns in Brahms Quartet No. 1 (1st movement)

rank	semitone interval pattern	# of instances	percent
1	+2 +1	44/660	6.7%
2	+3 +0	34/660	5.2%
3	+0 +3	28/660	4.2%
4	+2 -2	27/660	4.1%
5	-4 -3	26/660	3.9%
6	+4 +5	25/660	3.8%
7	+2 +2	22/660	3.3%
8	+4 +3	21/660	3.2%
9	+4 -4	19/660	2.9%
10	+3 +5	18/660	2.7%
11	+4 +1	15/660	2.3%
12	+5 +4	12/660	1.8%

Further inspection of Table 7 reveals that the dozen or so most prevalent patterns show a marked predominance for ascending intervals. This bias toward ascending pitch sequences is confirmed in Table 8. Table 8 tabulates the predominance of all *up-up, up-down, down-up,* and *down-down* pitch contours in opus 51, no. 1. Of 553 interval diads (pitch triads), nearly 40 percent are solely upward in their contours -- twice as many occurrences as the *down-down* contour. This suggests that the specific interval sizes may be less important in our feature definition than we might think. The right-most column of Table 8 shows the number of occurrences of the various contours -- where all of the alpha interval-class occurrences have been excluded. As can be seen, the *up-up* pattern continues to predominate. In other words, the *up-up* pattern remains prevalent in this movement independent of the alpha interval patterns.

Table 8

Pitch Contour Patterns for Brahms Quartet No. 1 (1st movement)

-----------	Instances	-----------
Contours	alpha included	alpha excluded
up-up	198 (37.1%)	151 (31.8%)
down-down	101 (18.9%)	90 (18.9%)
up-down	106 (19.9%)	106 (22.3%)
down-up	128 (24.0%)	128 (26.9%)
	553	475

Returning to Table 7, we can see that the second- and third-ranked patterns outline the minor third interval -- an interval that Forte also highlights in his analysis. Unfortunately, Forte's discussion of the relationship between alpha and the ascending minor third leaves the omission of the second pitch as a mysterious compositional act. By contrast, when viewed in the rhythmic-metric context, the elimination of the second pitch is far less enigmatic. Given the long-short-long contexts of the minor-third occurrences, it is clear that the `middle' note of (+2,+1) has been treated as a dispensable "unaccented passing tone" -- a status that befits a note that we now know is interposed between two notes of longer duration and which occurs in a weaker metric position than either of its neighbors.

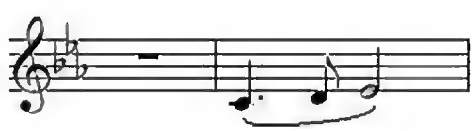
The sixth- and eighth-ranked patterns in Table 7 show what can only be interpreted as upward rising major and minor triads, whereas the seventh-ranked pattern corresponds to clear statements of the upward rising motive in the major mode (i.e., measures 232 to 260). However, due to the rejection of diatonic intervals in existing set theory, the (+2,+2) interval pattern must be defined as a separate set (Forte's epsilon motive, see Example 2).

With regard to diatonic intervals, a further analysis found that all of the instances of the prime form of alpha occuring in the first movement of quartet No. 1, are spelled as an ascending major second followed by an ascending minor second. By contrast, nearly ten percent of the alpha retrograde instances are spelled enharmonically as diminished thirds followed by minor seconds.

Feature Description

In light of the above analyses, we can now summarize the state of our alternative motivic description of the principal foreground feature in the first movement of Brahms's opus 51, no. 1. The feature consists of pitches in an upward-rising sequence, linked to a long-short-long rhythm, generally starting a phrase or slur, often following a rest, beginning in a strong metric position, and most likely to occur in an outer voice. (See Example 3.)

Example 3.



Schematic representation of the principal motive in Brahms Op. 51, No. 1, mov. 1 as developed using a comparative analytic method.

The feature bears more than a superficial resemblance to the opening statement in the first violin. In traditional tonal theory, this feature would have been informally labelled `the principal motive.' All of the above analysis has merely formalized the evidence in support of this informal intuition.

It is important not to draw the wrong conclusion from the foregoing analysis. One might suppose that the analysis reinforces three old criticisms of set theory, especially in the analysis of tonal music:

1. Set theory fails to account for rhythmic aspects of the music.
2. Set theory specifies intervalic sizes that are inappropriately fixed and overlooks the essentially diatonic contexts of the intervals. And
3. Set theory implies an equivalent or enhanced status for the retrograde, inversion, and retrograde inversion forms of the feature that are often not borne-out by the work itself.

According to the view presented in this paper however, the above criticisms are merely artifacts of a more fundamental problem. The nub of the problem with the alpha pattern is that it fails to define a feature that will distinguish the first movement of quartet No. 1 from the first movements of Brahms's other two string quartets. That is, the alpha pattern is akin to describing a robber as having a nose and two eyes. [13] [14]

In the case of Brahms's opus 51, no. 1, the above three criticisms are indeed warranted. But they are not warranted because rhythm is inherently important, diatonic intervals are privileged, or set theory exhibits a bias toward certain set variants. There is no telling how any given work may distinguish itself: rhythm may or may not be important, diatonic intervals may or may not be important, and set variants may or may not be important. In our analysis, these properties became important simply because they help to define a feature description that distinguishes the object under consideration from other similar objects.

Review of Analytic Assumptions

In presenting the above analysis, there are a number of caveats and disclaimers that must be made explicit. First, it is important to recognize that the above analysis dealt only with intratextural elements and did not consider possible intertextural features to opus 51, no. 1. It is always possible that intertextural properties overshadow the intra-textural elements. As noted earlier, a single statement of, say, "B-A-C-H" may prove to be of great significance, even though such a statement may not be prevalent. Forte proposes a plausible (and fascinating) intertextual property of this movement when he links the C minor/Eb minor key relations of the first and second subjects to those of the opening movement of Beethoven's opus 13 piano sonata (*Pathétique*). This observation is consistent with the view that Beethoven's work provided a model for Brahms's writing.

Second, only a handful of salience-enhancing properties were explored in the foregoing analysis. Specifically, we examined prevalence, metric position, voice position, primacy, and recency-linked properties. In the specific case of metric and primacy properties, there are alternative ways of measuring such properties -- other than those used in this paper. Other accent types (such as melodic accent), and mnemonic properties were not explored in the above analysis. In addition, other ways of characterizing features -- such as scale-degree, diatonic interval (rather than semitone intervals), or relative interval size (e.g. step/leap) -- were not explored. Rather than investigating long/short durational relations, a better rhythmic feature might arise from characterizing the precise durational proportions (e.g., the 3:1 ratio of the dotted rhythm). Any of these alternative approaches might lead to a better feature description than the one shown in Example 3.

Third, the above analysis was restricted to foreground features only. In principle, however, the same comparative methodology can be used to evaluate claims of background features. In the case of Schenkerian analysis, this approach must await some computable implementation -- if this is possible.

Fourth, the above claims are limited by the comparison group of works. The failure of the alpha pattern to distinguish Quartet No. 1 from Brahms's other string quartets does not mean it is not a distinctive feature of some larger group of objects. For example, the original alpha pattern might distinguish Brahms's string quartets from works by other composers, or other works by Brahms. Or it may be that the alpha pattern is characteristic of Brahms in general, or of Western tonal music compared to other types of music.

Once again it bears emphasizing that the motivic feature shown in Example 3 is not claimed to be the "best" intratextural feature in opus 51, no. 1. In principle, it is impossible to determine the best feature, since there may always be some analytic perspective that provides a more penetrating characterization. My claim is merely that our refined feature is better than Forte's original alpha motive.

Conclusion

This paper has presented an analytic method for refining and evaluating feature descriptions. The method is inherently comparative -- that is, it clarifies features by placing a work in the context of a *musical corpus*. The method highlights what makes one work different from another.

The method allows different musical parameters to be integrated within a unified feature description. For example, the method allows the analyst to resolve whether or not particular articulation marks are integral components of a motive.

In applying the method to Brahms's opus 51, no. 1, the above partial re-analysis leads to a number of conclusions:

First, it was found that although Forte's alpha interval-class pattern is prevalent (occurs frequently) in opus 51, no. 1, comparison with other string quartet movements by Brahms indicated that the pattern is not distinctive. Distinctiveness arose only when this pattern was linked to contexts deemed *a priori* to further increase their salience. This observation has a number of repercussions. It suggests that the concept of salience is not a mere fiction. It also suggests that our operationally-defined estimates of salience (such as phrase onset or outer-voice position) can be useful indices of salience. In our analysis, salience was measured in terms of phrase-related primacy and recency, rest-induced grouping, unison doubling, outer-voice position, and metric position. All of these factors have been shown experimentally to enhance the perceptibility of musical events. That is, at least some forms of salience arise from known perceptual phenomena. In short, our analysis provides additional evidence that perceptual factors are often important in musical analysis.

Second, we showed that it is not the alpha interval-class pattern that is the principal feature of this movement, but a particular form of the pattern. Although the concept of interval-class identity may be important in certain kinds of music, there is no indication in this movement by Brahms's that interval-class patterns play any significant role.

Third, we showed that the pitch-interval feature was inextricably linked to a particular rhythmic context -- namely, long-short-long. Specifically, it was demonstrated that this link is bi-directional: the pitch contour is most commonly associated with the rhythm, and instances of the rhythm are most commonly associated with the pitch contour. In extending set theory to tonal works, theorists ought to develop methods for characterizing pitch-rhythm interdependencies.

Finally, it should be noted that the concept of "distinctiveness" is simply an introduction to music analysis of one of the most important concepts in contemporary scientific method -- namely the idea of the "disproof of the null hypothesis." The basic idea is that whenever an observation is made, scholars should endeavor to adjudicate its pertinence by asking the question "How likely would it be that this observation might arise merely by chance?" In experimental methodology, it is the so-called "control group" that plays a critical role in establishing the likelihood of observing something by chance. In the foregoing analysis, we have shown that examining a comparison group (in this case the opening movements of Brahms's other two string quartets), permitted us to advance more readily in identifying a distinctive feature of opus 51, no. 1. [15]

Footnotes

[1] Forte's analysis of this work was first published in 1983 and subsequently reprinted in 1987. "Motivic design and structural level in the first movement of Brahms's String Quartet in C minor." *Musical Quarterly,* Vol. 69, No. 4 (1983 Fall) pp. 471-502. Reprinted in: Michael Musgrave (editor) *Brahms 2; Biographical, Documentary and Analytic Studies.* Cambridge: Cambridge University Press, 1987; pp. 165-196.

[2] See Forte's *The Structure of Atonal Music.* New Haven: Yale University Press, 1973; also "Bartok's `Serial' Composition." *Musical Quarterly,* Vol. 46, No. 2 (1960) pp. 233-245.

[3] On the contrary, I regard Forte's goal of increased rigor in analytic tasks as a goal worthy of the flattery of imitation.

[4] Further discussion regarding this point may be found in Leonard Meyer's *Emotion and Meaning in Music* Chicago: Chicago University Press, 1956.

[5] The assumption that lines-of-sound are psychological "real" rather than "reified" is supported by a wealth of perceptual research. As theorists are well aware, not all pitch successions evoke intervals. For an extensive review of the pertinent perceptual evidence see Albert Bregman, *Auditory Scene Analysis,* MIT Press, 1990.

[6] This observation is chronicled in detail in David Huron and Jonathon Berec, "A Method for characterizing instrumental idiomaticism: A case study of the B-flat valve trumpet." MS.

[7] See, for example, the discussion of quotation and allusion in Kenneth Hull, *Brahms the Allusive: Extra-compositional reference in the instrumental music of Johannes Brahms.* PhD dissertation, Princeton University, 1989.

[8] There are other possible definitions of a good description, although I am not aware of any in the field of music theory. The definition proposed here merely echoes the widespread notion in theory (promoted by Schoenberg) that music and musical descriptions ought to seek an economy of expression. It is noteworthy that this definition of good description is analogous to the concept of *efficiency* in technical disciplines.

[9] An introduction to the historical background of this work may be found in Michael Musgrave and Robert Pascall, "The String Quartets Op. 51 No. 1 in C minor and No. 2 in A minor; a preface." In Michael Musgrave (editor) *Brahms 2; Biographical, Documentary and Analytic Studies.* Cambridge: Cambridge University Press, 1987; pp. 137-143.

[10] All of the ensuing measurements were carried out using the Humdrum Toolkit software. All repeats were expanded in the electronic scores prior to processing. See David Huron, *Unix Software Tools for Music Research; The Humdrum Toolkit Reference Manual.* Menlo Park, CA: Center for Computer Assisted Research in the Humanities, 1995.

[11] A ludicrous comparison to be sure, but one that well illustrates the point.

[12] The term "significant" is used here in the formal statistical sense of the word. Pooling the data for the two control movements, a chi-square analysis for the ratios of expected to actual instances produces the following results: prime form of alpha (X=56.47; df=1; p<<0.001 significant); inverted form (X=1.82; df=1; p=0.18, not significant); retrograde form (X=53.39; df=1; p<<0.001, significant absence); retrograde inverted form (X=5.22; p=0.02, significant absence).

[13] Forte's motivic description would appear to have only one advantage over the feature developed in this paper and shown in Example 3: namely, the alpha motive is more succinct. However, since the alpha interval-class motive is not distinctive of the work in question, the feature description must be regarded as too brief.

[14] Readers familiar with Forte's paper may have noted that our analysis has focused exclusively on Forte's alpha motives without mentioning the innumerable other subsidiary patterns discussed in his analysis. The reason should now be clear. Many of Forte's other motivic sets suffer from the same problems, and other sets appear to be attempts to patch-up the short-comings of the alpha pattern.

[15] This research was undertaken while the author was visiting scholar at the Center for Computer Assisted Research in the Humanities, Stanford University. The author is grateful to the Center's Director, Dr. Walter Hewlett, for providing critical feedback and advice.

Scarlet & Grey
Ohio State University
School of Music

Musical Expectation

David Huron

Table of Contents

Abstract

Research pertaining to musical expectation is reviewed and an over-arching theoretical framework is described. Expectations originate in evolutionarily adaptive mechanisms for anticipating future events. Accurate expectations facilitate gathering information from the world and aid in preparing appropriate motor responses. Expectations are learned in response to statistical regularities evident in one or more cognitively encapsulated environments. Cognitive encapsulation permits the co-existence of different

"genres", so diverging expectations can arise depending on how the listener conceives of the genre. The statistical heuristics that drive expectations may vary in their accuracy and some heuristics may evoke expectations that exhibit systematic errors. Independent expectations relate to the "what" and "when" of possible outcomes. *Tonality* is one manifestation of "what"-related expectations, whereas *meter* is one manifestation of "when"-related expectations. Expectations may pertain to both immediate successions of events and to more distant contingent events. Expectations can arise from general-purpose schemata or from episodic memories. Sometimes these two memory systems predict different outcomes, accounting for such phenomena as the "surprise" of a deceptive cadence that is otherwise entirely expected. In addition, expectations can adapt dynamically as events unfold. Four sources of expectation-related emotion are distinguished: pre-outcome *imaginative* and *tension* responses, and post-outcome *appraisal of the outcome*, and *appraisal of the expectation*. Using these resources, musicians have become adept at crafting specific

emotional effects. Several musical examples are analyzed.

Table of Contents

Preface

Expectation is a constant part of mental life. A cook expects a broth to taste a certain way. A pedistrian expects traffic to move following a green light. A poker player expects an opponent to bluff. A pregnant woman expects to give birth. Even as you read this book, you have many unconscious expectations of how a written text should unfold. If my text were abruptly to change topics, or if the prose suddenly switched to a foreign language, you would naturally be dismayed. Even less dramatic changes would still have an effect. Some element of surprise would occur if a sentence proved ungrammatical, or if a sentence ended. Prematurely.

Half a centry ago, Leonard Meyer drew attention to the importance of expectation in the listener's experience of music. Meyer's seminal book, *Emotion and Meaning in Music,* argued that the principal emotional content of music arises through the composer's choreographing of expectation. Meyer noted that composers sometimes thwart the listener's expectation, sometime delay the expected outcome, or simply give the listener what is expected. *Emotion and Meaning in Music* was written at a time when there was little general experimental or theoretical psychological groundwork to draw upon. In the intervening decades, a considerable volume of research has accumulated. This research provides an opportunity to revisit Meyer's topic, and to recast the discussion in light of

contemporary findings. The principal purpose of this book is to fill-in the details and to describe a more comprehensive theory of musical expectation -- a theory I call the "ITPO" theory.

My motivation for developing the theory originated in purely musical ambitions. However, in piecing together a theory of musical expectation, it became clear that the ITPO theory amounts to a general psychological theory of expectation. Accordingly, psychologists are apt to find the theory of interest, even if they have no interest in music. While the theory itself will be described in general terms, the illustrations will be drawn almost entirely from the field of music. Psychologists may wish to skip much of the applied discussion that dominates the latter half of the book. Parallel examples in visual perception, linguisticis, social behavior, and ethology will readily come to mind for those readers who are knowledgeable in such areas.

In recent decades, much of the experimental research pertaining to expectation has focussed on auditory and musical expectation. This is unusual in the world of psychology, where research on perception is dominated by vision. Perhaps the inherently dynamic nature of sound encouraged a greater curiosity about the problem of expectation among auditory researchers. It may well be the case that the experimental research related to musical expectation represents the most advanced of any of the literatures pertaining to expectation. As a musicologist interested in the psychological experience of music, I am all too aware of how much my work has benefitted from discoveries in general psychology and cognitive science. So it is gratifying to imagine that music scholarship may be able to repay some of our debt to psychology.

I should at the outset confess that this book began in a uniquely inauspicious way. It began as a file-folder containing ideas that I didn't want to write about. In 1999, I was invited to give the Ernst Bloch lectures at the University of California, Berkeley. My lectures were entitled "Foundations of Cognitive Musicology", and one of the six lectures concerned a theory of music and emotions I had assembled. At the time, I thought that expectation was a comparatively minor component of the emotional experience of listeners. To be sure, I did not think expectation was unimportant. I just thought that other aspects of auditory-evoked emotion were more central. In writing up my theory of music and emotion for later publication, I bracketed "expectation" as a topic that would be explicitly excluded. Nevertheless, as expectation-related issues surfaced, I wrote brief summaries and tossed them into the folder labelled "Ignore This." As my work on music and emotion progressed, more and more slips of paper were relegated to this file.

In the spring of 2001 I taught a graduate seminar entitled "Music and Emotion". Once again I was eager to segregate the phenomenon of expectation from what I considered the main curriculum. Of course the topic could not simply be hidden from the inquiring minds of my students. I wrote a document entitled "Musical Expectation" whose sole purpose was to provide a stand-alone resource that students could read outside of class. I wanted to prevent the phenomenon of expectation from spilling into what I really wanted to talk about in class.

As I began to write the document, it became clear that my file on expectation had become the proverbial 500 pound gorilla in the filing cabinet. I finally recognized that I could

no longer ignore the role of expectation in musically-evoked emotion. The document expanded into an article, and finally into this book. In short, this book began as a "negative" endeavor -- a file of things I wanted to exclude from my other writing projects. Having admitted to its trash-pile origins, I sincerely hope that the finished product has transcended its inauspicious beginnings, and that readers will find that the theory contained here coherent and worthwhile.

A number of colleagues, collaborators, students, and friends have contributed directly or indirectly to this book. I wish to make explicit my debts of gratitude, and to express publicly my most sincere thanks. Much of this work was inspired by research carried out in my lab by post-doctoral fellow Paul von Hippel and graduate student Bret Aarden. Although their work was mostly carried out in my lab at the Ohio State University, I was not always a collaborator. Paul von Hippel took my suggestion about the possible influence of regression-to-the-mean in melodic organization and was able to make major discoveries that produced a silk purse from a sow's ear. In particular, von Hippel's experiments made clear the discrepancy between the reality and appearance of expectations. Bret Aarden took my interest in reaction-time measures in judging melodic intervals, and turned the paradigm into a truly useful tool for investigating musical expectation. His work has transformed the way we understand the Krumhansl and Kessler key profiles.

Although I have always preferred so-called "structural" theories of tonality to "functional" theories, I have benefitted enormously by having David Butler (the principal advocate of functional tonality) as a departmental

colleague. Professor Butler's persistent and knowledgeable criticisms of structural theories led me to better understand the importance of concurrent mental representations.

My discussion of rhythmic expectation builds directly on the research of my colleague, Prof. Mari Riess Jones working in the Ohio State University Department of Psychology. Along with her collaborator, Ed Large, they assembled a theory of rhythmic attending that has proved valuable for understanding the "when" of expectation.

Other colleagues provided stimulating conversation, correspondence, critiques and encouragement, including Caroline Palmer, Kristin Precoda, Simon Durrant, Don Gibson, Jonathan Berger, and Joy Ollen. To all of these individuals, my heartfelt thanks.

Finally, I am indebted to the Ohio State University. Of the various institutions I have worked at, this institution has an unusually high concentration of enlightened administrators. The university's support for music cognition has been visionary and unprecedented. Since the 1960s, the OSU School of Music has both tolerated and promoted the systematic and empirical study of music. I am grateful for the supportive and productive research environment.

Introduction

The theory of expectation proposed here is explicitly founded on principals of evolutionary psychology. From the evolutionary perspective, two questions help to frame the problem: (1) Why did the mental capacity to form expectations arise? That is, what are the adaptive purposes of expectations? And (2) Why might expectations evoke

various feeling states? That is, what are the adaptive purposes of the emotional responses that are conjured up due to expectations?

As a starting point, the ability to form accurate expectations about the future is clearly a potentially valuable biological function. It would be an advantage for an animal to be able to anticipate (say) that the trajectory of an approaching object is likely to intercept its path. Similarly, it would be an advantage for an animal to anticipate that eating food in a less conspicuous spot might reduce the likelihood of attracting a crowd of hungry competitors.

Of course it is possible that such behaviors might arise without any phenomenal sense of expectancy. For example, an animal might simply have an innate disposition to seek isolation when eating food. Similarly, the response to avoid an approaching object might merely arise as a conditioned reflex. In each case, it is possible that no "expectation" is involved. The animal need not "expect" that other animals might want to take its food, or "expect" that it will be struck by a moving object unless it takes evasive action. How do we know when a mental state might be properly regarded as one of "expectation?"

One defining characteristic is that the object of expectation is an event in time. Accordingly, an expectation entails some conscious or unconscious mental representation of such a future event. Consider again an animal moving to avoid collision with an approaching object. We might say that this is a conditioned response if the animal holds no mental representation of the hypothetical event of a collision. Similarly, consider again an animal moving food

to a less conspicuous location. If the animal takes this action based on some vague anxiety, or some antipathy about being observed, then the phenomenon is not properly speaking one related to expectation. However, if the animal forms a mental image or representation, say of other animals stealing its food, then we might regard the animal's actions as a proper consequence of a mental expectation.

In short, expectation, as conceived here, has something to do with the generation of mental representations of alternative possible future states. In this sense, expectation may be regarded as a cognitive phenomenon. These mental representations might be very complex and conscious, as when a politician imagines what will happen should she win an election. However, the research suggests that most of the mental representations related to expectation are unconsious and engage considerably simpler representations. From a behavioral perspective, it may be quite difficult to determine whether a particular animal behavior arises due to expectation or from some other process. I expect that the differences will become more apparent with advances in neurophysiology. Moreover, from an evolutionary perspective, there may be little difference whether an adaptive behavior is evoked due to a conditioned reflex, or whether it arises from a physiological process we might call "expectation."

The object of expectation is an event in time. Two principal types of uncertainty attend expectation: *what* will happen and *when* will it happen. And (2) Why might expectations evoke various feeling states? I propose that what we call "expectation" involves four functionally distinct physiological systems. Each of these systems can evoke

responses independently. The responses involve both physiological and psychological changes. Some of these changes are autonomic, and might entail changes of attention, arousal, and motor movement. Others involve noticeable psychological changes such as rumination and conscious evaluation. Outcomes matter, so positive and negative (i.e. "valenced") feeling states can also arise. It is the possibility of influencing these feeling states that attracts musicians to the phenomenon of expectation.

As it happens, these four systems tend to be invoked at different times. Consider Figure 1. Figure 1

Fig. 1. Schematic diagram of the "ITPO" theory of expectation. In expecting some future event, four response systems are activated successively. Feeling states are first activated by imagining different outcomes (I). As the anticipated event approaches, physiological arousal typically increases, often leading to a feeling of increasing tension (T). Once the event has happened, some feelings are immediately evoked related to whether one's predictions were born-out (P). Finally, feel states are evoked that are directly related to the value of the outcome (O). See text.

There are a number of issues related to expectation that will be addressed in this book. It is appropriate to take a moment to identify in advance some of these issues. One issue is the so-called Wittgenstein's paradox. Wittgenstein raised the problem of how it is possible to be surprised by something we know will happen. In music, an excellent

example of Wittgenstein's paradox can be found in the deceptive cadence. The deceptive cadence (usually a *V-vi* harmony) will continue to sound "deceptive" even in musical works that are total familiar to a listener. In Chapter X, we will show how this paradox is resolved by the distinction between viridical and schematic memory. Several other interesting listening phenomena will fall out of this distinction.

Another issue is how is it that we can coherently hold expectations for different genres. For example, we expect a classical string quartet not to exhibit syncopation, but we expect a jazz number to be syncopated. Are there cross-genre influences? Does a modern listener's experience with jazz lessen the effect of a hemiola when listening to a renaissance motet? How rapidly do listeners adapt when listening to a new genre, or new work? Are there piece-specific expectations? What is the relationship between expectation and pleasure?

Expectation

The world provides an endless stream of unfolding events that can surprise, delight, frighten, or bore. The capacity to form accurate expectations about future events confers significant biological advantages. Those who can predict the future are better prepared to take advantage of opportunities and sidestep dangers. Over the past 500 million years or so, natural selection has favored the development of perceptual and cognitive systems that allowed organisms to anticipate future events. Like other animals, humans come equipped with a variety of mental capacities that help us form expectations about what is

likely to happen. Anticipating future sounds is one of the evolved functions of the auditory system. This capacity to anticipate the course of acoustical events inevitably influences how listeners experience music. Moreover, musicians have learned how to manipulate such expectations in order to achieve specific types of responses.

Music scholars have long observed that listening to music engages a mental disposition to anticipate. Some theorists have made expectation a centerpiece in their music theorizing (e.g., Berger, 1990; Gjerdingen, 1988; Kramer, 1982; Larson, 1999; Lerdahl & Jackendoff, 1983; Meyer, 1956; Narmour, 1990, 1992). Many other theorists discuss expectation in the context of other musical phenomena (e.g., Aldwell & Schachter, 1989; Hindemith, 1944; Piston, 1978; Rameau, 1722; Riemann, 1903; Schenker, 1906). In the 1950s and 1960s, writings on musical expectation drew inspiration from the new field of information theory (e.g., Cohen, 1962; Coons & Kraehenbuehl, 1958; Kraehenbuehl & Coons, 1959; Moles, 1958/1966; Pinkerton, 1956; Youngblood, 1958). More recently, musical expectation has attracted the attention of experimentalists (e.g., Aarden, 2002; Abe & Oshino, 1990; Bharucha, 1994; Bigand & Pineau, 1997; Carlsen, 1981; Cuddy & Lunney, 1995; Dowling & Harwood, 1978; Federman, 1996; Francès, 1958; von Hippel, 1998; Jones, 1990; Jones & Boltz, 1989; Krumhansl, 1999; Rosner & Meyer, 1982; Schellenberg, 1997; Schmuckler, 1989; Sloboda, 1992; Thompson, Balkwill & Vernescu, 2000; Unyk, 1990; Werbik, 1969). At the same time, the general phenomenon of expectation has received sustained attention among psychologists working

in a number of diverse fields (e.g., Mandler, 1975; Olson, Roese & Zanna, 1996).

In the first instance, organisms would not be able to anticipate future events if the real world did not exhibit some structure. It would be impossible to predict an amorphous world that was devoid of any discernable patterns. Fortunately, the world exhibits many regularities. These regularities provide a useful starting point for understanding the nature of expectation. Expectations can be viewed as hypotheses about the structures underlying real world events (Shepard, 1981).

An important form of regularity is simple *event frequency* -- the tendency for some events to occur more frequently than others. You are more likely to hear the sound of a bird singing outside your window than the sound of a falling tree. You are more likely to hear a human voice than a bassoon. You are more likely to hear someone say "hello" than "hell no". And you are more likely to hear the pitch C4 than G#8. As we will see, music perception research has established that listening experiences are strongly shaped by such simple event frequencies.

Another form of regularity arises from the fact that some auditory events are contingent upon other events. The sound of my neighbor's car pulling into her driveway has a strong likelihood of being followed by the barking of her dog. The sound of a hammer striking a nail is likely to be followed by a repetition of the same sound. A dominant chord is more likely to be followed by a tonic chord than by a mediant chord. As with event frequencies, music perception research has established that *contingent frequencies* also influence the way music is experienced.

An additional aspect of expectation is the *environmental context*. The words a person is likely to utter can change dramatically depending on the situation. The words spoken by a robber holding a gun are predictably different from the words spoken by a man on his knees holding an engagement ring. We may anticipate that the sounds about to be emitted from a singer standing in front of an orchestra will differ from the sounds arising from a singer standing in front of a jazz trio. Expectations shift depending on such environmental contexts. Moreover, in order to form accurate contextual expectations, minds must learn to distinguish and recognize different contexts.

The fact that the world exhibits patterns or regularities does not necessarily mean that we are capable of taking advantage of these regularities. We may fail to decipher, recognize or learn the patterns that exist. For several years I failed to realize that shocks from static electricity are much more likely when I wear a certain jacket. Many listeners fail to learn that the second movement in a multi-movement work is likely to have a slow tempo. Most of the patterns that exist in the world go unrecognized. It is this profusion of unrecognized patterns that provides grist for the enterprise of science.

Even when we do learn to recognize a pattern, we may not recognize the *right* pattern. Consider, for example, the manner in which the pacific bull-frog anticipates a meal. During the Second World War, American soldiers stationed on pacific islands discovered an unusually maladaptive frog behavior. Soldiers discovered that if they rolled lead pellets from a shot-gun shell toward a bull-frog, the frog would immediately thrust its tongue forward and eat the pellet.

Curiously, the frog would do this repeatedly, never learning to avoid consumming the lead shot.

This is not at all a nice thing to do to a frog. But the phenomenon highlights an important fact about frog behavior. The frog has an instinct to eat anything that is small and black and moving. That is, the pattern "small-black-moving" causes the frog to anticipate a meal. In most circumstances, this instinctive behavior is beneficial for the frog. But in exceptional circumstances, the frog's disposition is utterly inept. For the pacific bull-frog, this behavior is instinctive, so the frog is incapable of learning a more nuanced behavior. While there are important advantages to instinctive behaviors, the case of the pacific bull-frog vividly demonstrates why learned behaviors can be superior to pre-wired instincts.

This raises the general question of whether expectations are learned or innate. As we will see, there are excellent reasons why auditory expectations would be predominantly learned. Since we know that learning can be incomplete or inaccurate, we might also expect to see evidence of "poorly learned" expectations in sound and music. In PART I, we will consider in detail evidence concerning the nature of auditory learning.

Expectations can differ with respect to their time-frame. Some expectations pertain to the flow of immediately successive events, as when your eyes move predictably along a line of text. Other expectations relate to longer time-frames, as when a person anticipates a surprise birthday party several days in advance. In music, contingent events occur in both the short-term succession of notes, and in longer-term expectations, such as an

impending cadence, an anticipated modulation, or the expectation of the ensuing song on a recorded album. In PART II, we will examine the different time frames that exist in anticipating future musical events.

In PART III we will note that minds are sensitive to the contexts of different worldly regularities. Important cognitive functions have evolved in order to ensure that these contexts are segregated from one another. We will note that these encapsulated contexts make it possible for different musical styles and genres to exist. Several different sets of auditory expectations can co-exist within the mind of a single listener.

In PART IV a comprehensive theory will be proposed whose purpose is to account for the observed psychological consequences linked to expectations. As we will see, accurate expectations are rewarded -- even when the predicted outcome is unpleasant. Four different types of responses will be distinguished; two types of responses precede the stimulus, and two further types of responses follow with the advent of the stimulus. The theory will be illustrated by analyzing several musical passages. The theory is not restricted to musical or auditory phenomena, however, and can be applied to any expectation-related behavior.

It is the capacity for expectations to evoke largely predictable emotional responses that makes the manipulation of psychological expectation such a compelling phenomenon for musicians. We will see that a number of common compositional techniques can be plausibly attributed to the manipulation of listener expectations. At the same time, we will see the importance

of enculturation in establishing a background of auditory expectations that make it possible to use specific musical devices.

Following a summary conclusion, we will identify as yet poorly understood aspects of the psychology of expectation, and point to future research possibilities.

I. AUDITORY LEARNING

The Baldwin Effect

Whether it is best for a behavior to be instinctive or learned depends in part on the stability of the environment. When an environment changes relatively rapidly it becomes difficult for an adaptive instinct to evolve. Biological examples of this phenomenon abound. For example, the most flavorful insect eaten by a species of salimander keeps changing color markings every decade or so. Rather than providing the salimander with an instinct to eat insects with a fixed coloration, a better adaptation would provide the salimander with the capacity to learn which color markings are indicative of a tasty food source.

The idea that evolution can account for the capacity to learn without invoking a Lamarckian notion of inherited learning was postulated in 1896 by James Baldwin. An *evolved capacity to learn* is consequently referred to as the Baldwin Effect (Baldwin, 1896, 1909).

Conceptually, auditory expectations might include both innate and learned components. A small number of aspects of human audition appear to be innate. For example, loud unexpected sounds will reliably evoke a startle response in all animals that have a sense of hearing. This response engenders a number of physiological changes that prepare the individual for possible defensive action -- such as increased heart rate and perspiration. Similarly, the orienting response is an innate reflex that causes listeners to direct their auditory gaze at unexpected sounds. This response produces physiological and neurophysiological changes that facilitate gathering further information from the environment (Lang, Simons & Balaban, 1997). However, apart from a handful of such reflexes, the extant research strongly implicates learning. This reliance on learning, in turn, implies that the auditory environment in which humans evolved was characterized by a high degree of acoustic variability. Like the salimander eyeing the color markings of an insect, humans could not necessarily count on a given sound to have a reliable or invariant "meaning."

The Baldwin effect holds important repercussions for our understanding of music's creative future. If learning plays the preeminent role in forming auditory expectations, then this suggests that musicians may have considerable latitude in creating a wide range of musics for which listeners may form appropriate expectations.

The Problem of Induction

Before we begin talking about auditory expectations, we should consider how auditory learning takes place. Learning from experience is regarded by philosophers as

the premiere example of inductive reasoning. Induction is the process by which some general principle is inferred from a finite set of observations or experiences.

The 18th-century Scottish philosopher, David Hume, recognized that there are serious difficulties with the method of induction. Hume noted that no amount of observation could ever resolve the truth of some general statement. For example, no matter how many white swans one observes, an observer would never be justified in concluding that *all swans are white.* Epistemologists agree that, in contrast to deductive reasoning, inductive reasoning is inherently fallible. From a purely logical point-of-view, it is not possible to infer the true principles underlying the world, solely from experience.

At first, the problem of induction would seem to make "knowledge" about the world impossible. Clearly, organisms do indeed learn from experience. The problem of induction merely places restrictions on this knowledge. Inductive knowledge must be contingent and fallible. Inductive knowledge is vague and adaptive, rather than precise and logical.

How, we might ask, has nature addressed the problem of induction? On what basis do organisms form generalized principles about the patterns of the world? It appears that nature approaches the problem in a manner quite similar to the methods of empirical science. Experiential learning appears to be statistical in nature. *Most swans are white* is good enough.

One of the most important discoveries in auditory learning has been that listeners are sensitive to the probabilities of

different sound events. Learning occurs for both *event frequencies* and *contingent frequencies*.

Event Frequencies

Both humans and animals are attuned to the frequency of occurrence for various stimuli in their environments. This sensitivity to probabilistic patterns is evident in auditory, visual and tactile stimuli, and has been observed in a number of species (see Hasher & Zacks, 1984; Gallistel, 1990; Kelly & Martin, 1994; Reber, 1993 - as cited in Saffran *et al,* 1999).

Perhaps the best example of event frequency learning in music is the phenomenon of absolute pitch. A person who possesses absolute pitch can name or identify the pitch of a tone without any external reference. Obviously, absolute pitch must involve learning since the pitch categories and labels are culture-specific. But the evidence for learning runs much deeper. People who have absolute pitch are slower at identifying some pitches than others. For example, the pitches C and G and more quickly identified than E and B; similarly, the pitches C# and F# are more quickly identified than D# and G# (Miyazaki, 1990; Takeuchi & Hulse, 1991). In general, identifying black notes is slower than white notes. Simpson and Huron (1994) carried out a study that simply tallied how often each pitch occurs in a large sample of music. As one might expect, white notes are more common than black notes, and pitches like C# and F# occur more frequently than pitches like D# and G#. Simpson and Huron went on to show that the relationship between speed of identification and frequency of occurrence follows a well-known law of

learning known as the Hick-Hyman law (Hick, 1952; Hyman, 1953). The learning occurs by simple exposure, and listeners learn best those sounds that have the highest event frequencies. Another way of interpreting the Hick-Hyman law is that perception is more efficient for expected stimuli than for unexpected stimuli.

First Impressions

Only a minority of listeners have the skill of perfect pitch. More commonly, listeners hear tones with respect to a scale context. In Western tonal music, pitches may tend to be heard as scale degrees.

If listeners have internalized a simple probability distribution of events based on past experience, then we might expect that listeners would tend to assume that the first thing they hear would correspond to the most common event. For example, since the tonic and dominant pitches are among the most common pitches in music, [1] we might expect listeners to assume that an isolated pitch will be the tonic or dominant. Conversely, we would expect that listeners might have difficulty hearing an isolated tone as an improbable scale degree. Recall that the purpose of expectation is to form accurate predictions about the world -- so it should come as no surprise that good listeners would tend to expect an isolated pitch to be the tonic.

In Huron (1999), musician listeners heard isolated tones and were asked to imagine the tone as a particular scale degree. For example, the pitch G#4 might be played and the listener instructed to imagine the tone as the dominant pitch. Once they were able to hear the pitch as the

specified scale degree, they responded by pressing a key. In order to ensure that listeners were responding honestly, a harmonic cadence was played immediately following the key-press, and listeners were asked to indicate whether the cadence corresponded to the correct key or not. Fig. 1 shows the average response times for only those responses where the listener correctly recognized that the cadence passage was in/out of the correct key.

Figure 1

Fig. 1. Average response times for listeners to hear an isolated tone as a specified scale degree. Data are shown only for responses where the listener correctly recognized that an ensuing cadence passage was in/out of the correct key.

As can be seen, the fastest average response time is for the tonic pitch, followed by the dominant. That is, listeners were most easily able to imagine an isolated tone as the tonic or dominant. Some scale tones, like the supertonic and subdominant, are somewhat slower. The especially slow processing for "fah" will strike musicians as odd, since it is not a notably rare pitch. However, if we look at the *initial notes* in a large sample of major-key melodies, it turns out that "fah" occurs least frequently of all the scale tones. Melodies tend not to begin with "fah", and this fact is reflected in the difficult listeners have in conceiving an isolated tone as "fah".

In effect, listeners tend to hear an isolated pitch as though it is the starting pitch of a major-key melody: listeners tend to form expectations that approximate the distribution

of melody-initiating tones. The most frequently occuring starting scale degrees prove to be the easiest to process mentally. [N.B. von Hippel has collected data about similarly echoes listeners assumptions of absolute pitch height]. Even before the first note of music is sounded, listeners have expectations. Moreover, once the first note sounds, listeners are already "jumping to conclusions."

For musicians, these experimental observations simply affirm our informal subjective intuition that listeners tend to assume that an isolated pitch corresponds to the tonic.

Contingent Frequencies

Event frequencies pertain to the simple likelihood of individual events without regard to preceding events. But humans and other animals also learn to anticipate sounds on the basis of what has just been heard. For example, the probability of hearing the tonic pitch is increased if we are currently hearing the leading-tone. These context-related regularities are referred to as *contingent frequencies* or *conditional probabilities*.

Jenny Saffran, Richard Aslin and their colleagues carried out a set of seminal experiments that demonstrate the statistical manner by which tone sequences are learned by listeners. Saffran, Johnson, Aslin and Newport (1999) constructed various musical "vocabularies" consisting of 3-note "figures." An example of a vocabulary consisting of six basic melodic figures is notated Fig. 2.

Figure 2

Fig. 2. Sample of six melodic figures used in Saffran *et al* (1999). Exposure tone sequences were constructed by randomly stringing together such figures.

Using these figures, Saffran *et al* constructed a long (seven minute) tone sequence that consisted of a random selection of the six figures. Fig. 3 shows a sample excerpt from the sequence; it begins with figure #2, followed by figure #4, followed by figure #6, followed by figure #5, and so on. The random sequences were constrained so that no individual 3-note figure was repeated twice in succession.

Figure 3

Fig. 3. Sample tone sequence used in the exposure phase of Saffran *et al* (1999). Sequences were constructed from the three-note figures shown in Figure 2. Tone sequences were constrained so no single figure was repeated twice in succession.

Twenty-four listeners heard the seven-minute sequence three times for a total of 21 minutes of exposure. Note that the listeners had no prior knowledge that the tone sequence was conceptually constructed using a vocabulary of 3-note figures: listeners were simply exposed to a continuous succession of tones for 21 minutes.

In order to determine whether listeners had passively learned to preferentially recognize any of the 3-note figures, the 21-minute exposure phase was followed by a test phase. For each of 36 trials, listeners heard two 3-note stimuli. One stimulus was selected from the six vocabulary

items whereas the other 3-note stimulus had never occurred in the entire tone sequence. A sample test item is illustrated in Fig. 4 -- the first sequence is a vocabulary item whereas the second sequence is not:

Figure 4

Fig. 4. Sample test stimuli used in Saffran *et al* (1999). Listeners heard two three-note sequences and were asked to identify which sequence was more familiar.

Listeners were asked to identify which of the two 3-note items was more familiar. The results were clear: listeners correctly identified the three-note sequences they had been exposed to.

A possible objection to Saffran's experiment is that 4 out of 6 of the vocabulary items end on the pitches of a D major triad (D, F#, A). The pitches used in this experiment are consistent with the key of D major, so perhaps Saffran's listeners were merely preferring test items that implied some tonal closure.

Actually, the experiment was a little more sophisticated. The twenty-four listeners were divided into two groups. Only half of the listeners were exposed to the tone sequences described above. The other listeners were exposed to a different sequence constructed from six entirely different vocabulary "figures." Both groups of listeners were tested, however, using precisely the same test materials. The pairs of three-note figures were organized so that what was a vocabulary item for Group #1 was a non-vocabulary item for Group #2 and *vice versa.*

What one group of listeners deemed "familiar" was the precise opposite of what the other group deemed "familiar."

This experimental control allows us to conclude that what listeners heard as a "figure" had nothing to do with the structure of the figures themselves, and relates only to their simple probability of occurrence. A simple linguistic analogy might help to clarify the results. Suppose you heard a long sequence of repeated syllables *... ababababababa ...* How would you know whether you were supposed to hear *ab, ab, ab, ab, ab* ... or *ba, ba, ba, ba, ba* ...? In effect, Saffran trained two different groups of listeners, one to hear the sequence as *ab, ab, ab ...* and the other to hear the sequence as *ba, ba, ba.* (In fact, in an earlier experiment, Saffran, Newport and Aslin (1996) had done exactly this for spoken syllables.) For each item in the test phase, one group of listeners heard as a figure what the other group heard as a non-figure and *vice versa.*

Saffran and her colleagues went on to repeat both experiments with 8-month old infants. Infants tend to stare longer in the direction of novel stimuli. By tracking head movements in the test phase, they were able to show that the unfamiliar figures were perceived as exhibiting greater novelty for the infants. Once again, the infants were divided into two groups and exposed to different random sequences. That is, in the test phase, what was a "vocabulary" item for one group of infants was a "non-vocabulary" item for the other group, and *vice versa*. In short, both infants and adults learned to recognize the most frequently occurring patterns -- whether tone sequences or phoneme sequences. Moreover, those patterns that occurred most frequently, were the patterns that both adults and infants best recognized.

It is important to note that there were no silent periods, dynamic stresses or other cues to help listeners parse the figures. From the listener's perspective, the figures might have consisted of 2-note groups, 3-notes groups, or some other group size or mixture of group sizes. Also recall that none of the figures were repeated twice in succession. Since two groups of listeners learned diametrically opposite "motivic vocabularies", the internal structure of the figures had no effect on the perception of grouping. This means that the only possible conclusion is that listeners were cuing on the simple statistical properties of various tone sequences. More precisely, listeners were learning the contingent frequencies: given pitch *X*, the probability of pitch *Y* is high, but the probability of pitch *Z* is low, etc.

The 21-minute period of exposure allowed listeners to form a sense of the likelihood of different pitch successions. Table 1 shows the long-term conditional probabilities for sequences using the six figures shown in Fig. 2. The vertical axis indicates the antecedent state (initial note) and the horizontal axis indicates the consequence state (following note). For example, the probability of the pitch `C' being followed by a `C#' is 0.056. That is, 5.6 percent of C's are followed by C#'s. By contrast, the pitch `C#' is never followed by the pitch `C'.

Table 1

	consequent state										
	c	c#	d	d#	e	f	f#	g	g#	a	b
c	0	0.056	0	0	0	0.056	0.056	0	0	0	0
c#	0	0	0.056	0	0	0	0	0	0	0	0
d	0.011	0	0.022	0.011	0	0.078	0	0.022	0	0.022	0.056
d#	0	0	0	0	0.056	0	0	0	0	0	0
e	0.011	0	0.011	0.011	0	0.011	0	0.011	0	0.011	0

f	0.056	0	0	0	0.056	0	0	0	0	0	0
f#	0.011	0	0.011	0.011	0	0	0	0.011	0	0.011	0
g	0	0	0	0	0	0	0	0	0.056	0	0
g#	0	0	0	0	0	0	0	0	0	0.056	0
a	0.011	0	0.067	0.011	0	0.011	0	0	0	0.011	0
b	0.011	0	0.011	0.011	0	0.011	0	0.011	0	0	0

Applying these probabilities to the original exposure sequence, we can identify the likelihood of each pitch-to-pitch transition. Fig. 5 provides a schematic illustration of the transitional probabilities for the sequence shown in Fig. 3. Thick lines indicate pitch successions that have a strong probability of occurrence. Thin lines are less strong. No line indicates a weak likelihood. Notice how the 3-note structure of the figures can arise simply by recognizing strong conditional probabilities. Indeed Saffran's experiments establish precisely this fact: in order for a listener to learn to hear this sequence as constructed from 3-note vocabulary "motives" the listener would have to recognize, in some sense, that the boundaries between vocabulary motives have relatively low probabilities.

Figure 5

Fig. 5. Sample exposure stimuli showing the long-term statistical probabilities of note-to-note transitions. Thick lines indicate high probability. Thin lines indicate medium probability. Absence of line indicates low probability.

The work pioneered by Richard Aslin and Jenny Saffran provides just one of many examples showing how people (and animals) learn from exposure. Much of the research in this area pertains to vision, but Saffran and Aslin have

shown that the same statistical learning processes occur for adult and infant listeners -- both when listening to speech as well as when listening to tone sequences. In effect, both adult and infant listeners build a representation of the transitional probabilities between adjacent tones in a tone stream, grouping together tones with high transitional probabilities, and forming figure-boundaries at locations in the tone stream where transitional probabilities are low. The statistical properties of the sequence are learned as a by-product of simple exposure, without any conscious awareness by the listener.

STATISTICAL PROPERTIES OF MUSIC

The work of Jenny Saffran and others has established that listeners are sensitive to the probabilities of different sorts of events. But in Saffran's work, the tone sequences exhibited properties that were based on purely artificial probabilities constructed for her experiments. If we want to understand music-related expectations then we should focus on whatever statistical regularities real music exhibits.

There are indeed a number of stable probabilistic relationships that can be observed in music. Some of these probabilities reflect properties of individual musical works. Huron (2001a) for example, has shown how comparative probabilistic analyses can be used to identify thematic and motivic features in a musical work and distinguish one piece from another. Other probabilities appear to reflect properties of particular styles or genres (Moles,

1958/1966). Yet other probabilities appear to reflect properties of music as a whole. We might begin our musical story by looking for statistical regularities that seem to characterize Western music in general.

Mental Representations

Before continuing we might ask what is it that listeners represent when they form mental analogs of probability structures? For example, are tone sequences represented as pitches or as intervals? Saffran's experiments do not address this issue. A variant of Saffran's experiments might present the test materials transposed upward or downward and compare the associated recognition scores with those for the untransposed materials. If there is no difference, then the result would suggest that listeners employ a *relative-pitch* or *interval* based mental representation rather than an *absolute pitch* based representation. Conversely, if transposed figures evoke only chance recognition, then the results would suggest that listeners rely on an absolute pitch-related representation.

So what *are* the mental representations used by listeners? Theoretically, possible representations might include absolute pitch, pitch chromas (or pitch classes), intervals, scale degrees, contours, duration, relative duration, metric position, harmonic functions, chord qualities, spectral centroids, or other concepts.

Experimental evidence suggests that all of these representations are used by at least some listeners in some listening situations. Clearly, absolute pitch representations are available only to a minority of listeners -- those with perfect pitch. Musical coding may involve several

concurrent representations; Dowling (1978), for example, has proposed that for melodies, the most important pitch-related representations are scale degree and contour. Despite the research, little is known at the moment about the mental coding of music.

In some circumstances, knowledge of the precise nature of the mental representation may not be important. A useful way to illustrate this is provided by information theory. The field of information theory (Shannon, 1948; Shannon & Weaver, 1949) has provided useful mathematical techniques for characterizing the probablistic relationships between events. Information theory inspired a number of music theorists throughout the 1950s and early 1960s. However, it was abandoned (for reasons that are not entirely clear) by about the mid 1960s. [2] Information theory provides a way to measure contingent probabilities. When rolling dice, for example, we know that the number rolled is independent of numbers previously rolled (this is true even for loaded dice). By contrast, other events exhibit contingent effects as when the occurrence of the letter "u" in English text is considerably increased when preceded by the letter "q".

Figure 6 plots the flow of information for the tune *Pop Goes the Weasel.* Information is plotted (in bits) for five different representations. For example, the upper-most plot shows information according to the probabilities of different scale degrees. The probabilities used in Fig. 6 were derived from an analysis of roughly 6,000 Western European folk songs.

Figure 6

Fig. 6. Information theoretic analysis of *Pop Goes the Weasel* showing changes of information (in bits) as the piece unfolds. Plotted information includes scale degree, scale degree succession (degree diad), metric position, melodic interval, and melodic interval succession (interval diad).

Notice that the information for both scale degree and melodic interval representations peak at the word "pop". For scale degree diad and interval diad the word "pop" coincides with the second highest information value -- with the maximum value following immediately after the word "pop". There appears to be an element of musical "surprise" at this point that is echoed in the lyrics. As a children's action song, this point is usually accompanied by some abrupt action, also suggestive of surprise.

Note, however, that there is no comparable information peak for metric position. That is, the interval/pitch/scale-degree may be relatively surprising, but the *moment* of its occurrence is not surprising. This highlights a distinction that can be made between the *what* and *when* of surprise. In some musical situations, the "what" is expected, whereas the "when" may be relatively unexpected. A well-known example is evident in the popular "Ode to Joy" from Beethoven's Ninth Symphony, where one of the phrases begins a beat early.

With the exception of the metric position information, all of the pitch-related information values are positively correlated. Table 2 shows a correlation matrix for the information content (measured in bits) for the various representations used in the above analysis of *Pop Goes the Weasel*. An analysis of a sample of 200 melodies from

American, Chinese, Dutch, Pawnee, and Xhosa sources confirms that these positive correlations are endemic.

Table 2

	degree	degree dyad	metric position	interval	interval dyad
degree	+1.00				
degree dyad	+0.45	+1.00			
metric position	-0.31	-0.05	+1.00		
interval	+0.17	+0.74	-0.00	+1.00	
interval dyad	+0.30	+0.90	+0.02	+0.77	

The fact that different musical representations are positively correlated is both an advantage and a disadvantage. The advantage is that it implies that we can proceed with a probabilitistic analysis of music with relatively little concern over the choice of representation. On the other hand, this high correlation invites onerous mistakes of interpretation (as we will see). Results of perceptual experiments may very well be consistent with a particular representation, but the same results are likely to be consistent with several other alternative representations as well. For example, a result that is consistent with small interval sizes, will also be consistent with successions of neighboring pitches, or with close pitch chromas, or with small log-frequency differences between fundamentals, or with small differences in spectral centroid, or with small critical band distances, or with tonotopic proximity along the cochlear partition.

Pitch Proximity

One of the best generalizations we can make about melodies is that they typically employ sequences of tones that are close to one another in pitch. This tendency to use

small intervals has been observed over the decades by innumerable researchers, including Ortmann (1926), Merriam, Whinery and Fred (1956), and Dowling (1967). Fig. 7 reproduces results in Huron (2001b) showing the distribution of interval sizes using samples of music from a number of cultures: American, Chinese, English, German, Hasidic, Japanese, and sub-saharan African (Pondo, Venda, Xhosa, and Zulu). For a broad range of cultures, the preponderance of intervals tend to be small. Only pseudo-polyphonic melodies (such as yodelling) fail to consist predominantly of small pitch movements.

Figure 7

Fig. 7. Frequency of occurrence of melodic intervals in notated sources for folk and popular melodies from ten cultures (n=181). African sample includes Pondo, Venda, Xhosa, and Zulu works. N.B. Interval sizes only roughly correspond to equally-tempered semitones.

In 1981, James Carlsen carried out an experiment to determine whether listeners tend to *expect* small interval continuations. Carlsen tested listeners from three different Western cultures: American, German, and Hungarian listeners. Although there were some differences between these three groups, all listeners showed a marked expectation for continuations involving small pitch movements.

Unlike Saffran, Carlsen's work did not explicitly establish that the expectation for small intervals is learned by exposure to the music. (It is theoretically possible that

these expectations might have some other origin.) But given the facts that melodies tend to use mostly small intervals, and that the auditory system is sensitive to frequently occurring phenomena, it is not unreasonable to suppose that listeners might have learned to expect small intervals. At a minimum, we can conclude that small pitch intervals are a common feature of real music, and that listeners appear to expect small intervals.

Step Inertia

Another property of melodic expectation pertains to what Paul von Hippel has called *step inertia.* This is the idea that small pitch intervals (1 or 2 semitones) tend to be followed by pitches that continue in the same direction. Music theorist Eugene Narmour has suggested that listeners form these sorts of "step inertia" expectations for melodies and has even suggested that these expectations might be based on innate dispositions (Narmour, 1990).

The first question to ask is whether melodies themselves are indeed organized according to step inertia. Is it the case that most small pitch intervals tend to be followed by pitch contours that continue in the same direction? The answer to this question is a qualified yes. Von Hippel examined a large sample of melodies from a broad sample of different cultures. He found that only descending steps tend to be followed by a continuation in the descending pitch direction. Roughly 70% of descending steps are followed by another descending interval. In the case of ascending steps, no trend is evident. Following an ascending step, melodies are as likely to go down as to continue ascending (see Table 3).

Table 3

	Followed by Ascending Step	Followed by Descending Step
Initial Descending Step	**30%**	**70%**
Initial Ascending Step	**51%**	**49%**

Probabilities for Step-Step movements in a large sample of Western and Non-Western musics

But what about listeners' expectations? Do listeners expect a step movement to be followed by a pitch movement in the same direction? Von Hippel (2001) carried out the pertinent experiment and measured listeners' expectations in a variety of melodic circumstances. Von Hippel's listeners heard a twelve-note sequence and were then asked to indicate whether they expected the next note to be higher or lower. The results showed that listeners do indeed expect descending steps to be followed by another descending interval. Surprisingly, listeners also expect ascending steps to be followed by another ascending interval. That is, the results are consistent with Narmour's suggestion of step inertia.

But real melodies exhibit a tendency for step inertia *only* for descending intervals. So why do listeners expect step inertia for both ascending and descending contexts? Von Hippel suggested a plausible logic as to why listeners "over-generalize" in forming their melodic expectations: Notice that since ascending steps have a 50-50 chance of going in either direction, there is no penalty for (wrongly) assuming that ascending steps should typically continue to go up. That is, the expectation for step inertia is no worse than chance for ascending contours. Since the strategy of

expecting step inertia pays off for descending intervals, listeners who form a step-inertia expectation will still, on average, have more accurate expectations than a listener who has no step-inertia expectation.

The "step-inertia" strategy is favored for another reason as well. Working at the University of Nijmegen in the Netherlands, Piet Vos and Jim Troost (1989) discovered that large melodic intervals are more likely to ascend and that small melodic intervals are more likely to descend. Fig. 7 shows the frequency of occurrence of ascending intervals for different interval sizes. The dark bars show the results for Western classical music whereas the light bars show the results for mainly Western folk music. Fewer than 50% of small intervals ascend. The reverse holds for large intervals:

Figure 7

> **Fig. 7.** Frequency of occurrence of non-unison ascending intervals. Dark bars: sample of 13 Western composers. Light bars: sample of Albanian, Bulgarian, Iberian, Irish, Macedonian, Norwegian, and American Negro folk songs. (After Vos & Troost, 1989.)

Since ascending steps occur less frequently than descending steps, there is even less of a penalty for wrongly expecting that an ascending step is likely to continue in the same direction. The bias favoring descending steps therefore further increases the likelihood that a step-inertia expectation will pay off.

There is one noteworthy complication that arises from Von Hippel's experiment. Von Hippel tested both musician and non-musician listeners. He found step-inertia expectations only for the musician participants. The non-musicians had no discernable pattern related to step-interval antecedents. It is plausible that musicians have more experience listening to music than non-musicians. If so, it may be that the origin of the step-inertia expectation is attributable to passive learning through extensive exposure.

Post-skip Reversal

We have seen that listeners expect melodies to consist mostly of small pitch intervals. Experienced listeners also expect that small intervals tend to be followed by pitches that preserve the melodic direction -- although musical melodies only exhibit step-inertia for descending intervals. What about expectations following large intervals?

For hundreds of years, music theorists have observed that large intervals tend to be followed by a change of direction. More specifically, most of the theorists who have commented on this purported phenomenon have suggested that large intervals tend to be followed by step motion in the opposite direction. Since most pitch intervals are small, any interval should tend to be followed by step motion. The important part of the claim is the idea that large leaps should be followed by a change of direction. Following Paul von Hippel, we can call this purported tendency *post-skip reversal* (von Hippel, 1998).

Once again, the first question to ask is whether actual melodies conform to this principle. Do most large leaps tend to be followed by pitches that change direction? In

1924, Henry Watt tested this idea by looking at melodic intervals in musical samples from two different cultures: Lieder by Franz Schubert and Ojibway Indian songs. Watt's results for Schubert are shown in Fig. 8.

Figure 8

> **Fig. 8.** Watt's (1924) analysis of intervals in Schubert Lieder. Larger intervals are more likely to be followed by a change of melodic direction than small intervals. Watt obtained similar results for Ojibway Indian songs. No data point corresponds to 11 semitone intervals because of the absence of such intervals in Watt's sample. From von Hippel and Huron (2000).

For intervals consisting of 1 or 2 semitones, roughly 25 to 30 percent of contours change direction. That is, the majority of small intervals continue in the same direction. However, as the interval size increases, the graph tends to rise upward to the right. For octave (12 semitone) intervals, roughly 70 percent of intervals are followed by a change of direction. (There is no data point corresponding to 11 semitones because there were no 11-semitone intervals in Watt's sample.) Watt found similar results for the Ojibway songs.

Von Hippel and Huron (2000) carried out further tests of this idea using a broader and more diverse sample of melodies from cultures spanning four continents: traditional European folksongs, Chinese folksongs, South African folksongs and Native American songs. Once again,

for each of these repertories, the majority of large intervals are indeed followed by a change of direction.

Von Hippel and Huron proposed a rather unexciting reason for the existence of post-skip reversal. Most large intervals tend to take the melody toward the extremes of the melody's range. For example, a large ascending leap has a good probability of placing the melody in the upper region of the tessitura or range. Having landed near the upper boundary, a melody has little choice but to go down. That is, most of the usable pitches lie below the current pitch. Similarly, most large descending leaps will tend to move the melody near the lower part of the range, so the melody is more likely to ascend than to continue descending.

Melodies do not simply wander around the range of human hearing by taking mostly small steps. Instead, melodies exhibit pitch distributions that show a central tendency. That is, melodies display a stable tessitura or range. The most frequently occurring pitches in a melody lie near the center of the melody's range. Pitches near the extremes of the range occur less commonly.

Statisticians have shown that whenever a distribution exhibits a central tendency, successive values tend to "regress toward the mean." That is, when an extreme value is encountered, the ensuing value is likely to be closer to the mean or average value. Regression-to-the-mean should not be regarded as a "phenomenon." There is no "force" or "magnet" drawing values toward the mean. Regression-to-the-mean is simply an artifact of the fact that most values lie near the center of the distribution.

When you encounter a tall person, the next person you encounter is likely to be shorter. But the shorter person is not "caused" by the previous encounter with a tall person. It is simply a consequence of the fact that most people are near average height. Similarly, when we encounter a high pitch, we must be careful about assuming that movement toward the high pitch will somehow "cause" the next pitch to be lower.

If post-skip reversal is a consequence of regression-to-the-mean, then we ought to see a difference for leaps, depending on where they occur in the range. Consider the ascending intervals shown in Fig. 9. In this schematic illlustration, the mean or median pitch for the melody is represented by the bold center line in the staff. The first ascending leap takes the contour above the median. Both regression-to-the-mean and post-skip reversal would predict a change of direction to follow. In the second case, the ascending leap straddles the median pitch. Once again, both regression-to-the-mean and post-skip reversal predict a change of direction. In the third and fourth cases, the two theories make different predictions. In the third case, the leap lands directly on the median pitch. Post-skip reversal continues to predict a change of direction, whereas regression-to-the-mean predicts that either direction is equally likely. Finally, in the fourth case, the leap lands below the median pitch. Here regression-to-the-mean predicts that the contour should continue in the same direction (toward the mean), whereas post-skip reversal continues to predict a change of direction. So how are real melodies organized? Are they organized according to post-skip reversal? Or according to regression-to-the-mean?

Figure 9

Fig. 9. Four hypothetical interval relationships relative to the median (or average) pitch (represented by the bold central line): (1) median-departing leap, (2) median-crossing leap, (3) median-landing leap, and (4) median-approaching leap. See also Figure 10.

In order to answer this question, von Hippel and Huron (2000) studied several hundred melodies from different cultures and different periods. For each melody we calculated the median pitch and we then examined what happens following large leaps. Our results are plotted in Fig. 10, for the case where a `skip' is defined as intervals larger than 2 semitones. The black bars indicate instances where an interval is followed by a change of direction. The grey bars indicate instances where an interval is followed by a continuation in the same direction.

Figure 10

Fig. 10. Number of instances of various melodic leaps found in a cross-cultural sample of music. Most large intervals that approach the median pitch continue in the same melodic direction. Large intervals that land on the median pitch are as likely to continue in the same direction as to reverse direction. Results support the phenomenon of melodic regression, and fail to support post-leap reversal.

If post-skip reversal is the important organizing principle, then we would expect to see taller black bars than grey bars in each of the four conditions. By contrast, consider regression-to-the-mean. This would predict that black bars should be taller than grey bars for the median-departing and median-crossing conditions (which is the case). For skips that land on the median pitch, regression-to-the-mean would predict roughly equivalent numbers of continuations and reversals (that is, we would expect the black and grey bars to be roughly the same height -- which is the case). Finally, in the case of median-approaching skips, regression-to-the-mean would predict that melodies ought to be more likely to continue in the same direction toward the mean (that is, we would expect the grey bar to be taller than the black bar -- which is again the case).

Von Hippel and Huron carried out further statistical analyses which reinforce the above result. With regard to large intervals, melodies behave according to regression-to-the-mean and are not consistent at all with the idea of post-skip reversal. The further the leap takes the melody away from the mean pitch, the greater the likelihood that the next pitch will be closer to the mean. If a leap takes the melody toward the mean, then the likelihood is that the melody will continue in the same direction. Incidentally, we tried a number of different definitions of "large" leap. The results are the same no matter how a leap is defined in terms of size. We also looked for possible "delayed" resolutions. That is, we looked to see whether the second or third note following a large leap tended to change direction. Once again, the aggregate results always conformed to regression-to-the-mean, but not post-skip reversal. This was true in Schubert, in European folksongs,

in Chinese folksongs, in sub-Saharan African songs, and in traditional Native American songs.

It bears reminding that most large intervals are indeed followed by a change of direction. (For skips of 3 semitones or greater, roughly two-thirds are followed by a reversal of contour.) But this is only because most large intervals tend to take the melody away from, rather than toward, the mean pitch for the melody.

Having investigated the organization of actual melodies, we might now turn to the question of what listeners expect. Even if melodies are not organized according to post-skip reversals, might it not be the case that listeners *expect* large intervals to be followed by a change of direction? Or do listeners expect the next pitch to move in the direction of the mean?

Once again consider our earlier analogy to people's heights. When we encounter a tall person, do we (1) expect the next person to be of average height (the "real" phenomenon) or (2) expect the next person to be shorter -- an artifact of (1)? This question was answered experimentally by Paul von Hippel (in preparation). Von Hippel played large intervals in a variety of melodic circumstances, and asked listeners to predict whether the melody would subsequently ascend or descend.

The melodic contexts were arranged so that some large intervals approached the mean and other large intervals departed from the mean. If listeners' expectations are shaped by post-skip reversal, then they ought to expect all large intervals to be followed by a change of direction. However, if listeners' expectations are shaped by

regression to the mean, then they ought to respond according to the register of the interval: intervals in the low register (whether ascending or descending) should be followed by higher pitches while high register intervals (whether ascending or descending) should be followed by a lower pitch.

The results were clear: the register or tessitura of the interval doesn't matter -- listeners typically expect large intervals to be followed by a change of direction without regard to the location of the median pitch. That is, listeners expectations follow the post-skip reversal principle, rather than regression-to-the-mean.

As before, these results apply only in the case of musician listeners. Von Hippel's non-musician listeners showed no systematic pattern of responses. This difference between musicians and non-musicians once again implicates learning.

II. REALITY VERSUS APPEARANCES

We have seen two examples where experienced listeners have established an expectation strategy that works in most circumstances, but is only an imperfect approximation of the actual structure of the melodies. By way of summary, we can now compare and contrast how melodies are *actually* structured with how experienced listeners *think* they are structured.

Actual Melodic Structure - Expected Melodic Structure

Melodies show the following organizational elements:

1. **Pitch Proximity.** Successive pitches tend to be near to one another. Pitch proximity is not merely an artifact of central tendency. That is, pitch proximity doesn't arise simply because most of the pitches in a melody lie near the center of the distribution. If pitch proximity were the only organizing principle for melodies, then melodies might look something like the pitch sequence shown in Fig. 11. Here we see a randomly generated "melody" in which the only constraint is a bias toward smaller rather than larger intervals. The result is a so-called "random walk" -- what engineers call *Brownian noise.*

   Recall that correct expectations ought to better prepare an organism -- either for appropriate action or for more efficient perception. In the case of pitch proximity, Deutsch (1978) showed that listeners are more efficient when processing tones preceded by small intervals than by large intervals. Similarly, Boomsliter and Creel (1979) found that when exposed to short tones, listeners are faster to form pitch perceptions when the stimuli are embedded in music-like sequences. By contrast, unprepared listeners take longer to form appropriate pitch sensations.

   Figure 11

> **Fig. 11.** "Brownian" or "random walk" melody. Successive pitches are constrained only by the principle of small distances to the preceding pitch.

2. **Central Pitch Tendency**. If real melodies were constrained only by pitch proximity, then long melodies would inevitably wander out of range at some point. However, like the vast majority of other phenomena in the world, the most frequently occurring pitches in melodies tend to lie near the center of some distribution. If a central tendency were the only organizing principle then melodies might look something like the pitch sequence shown in Fig. 12. Here we see a randomly generated "melody" whose distribution corresponds to a normal distribution, centered in the middle of the staff. Engineers call this kind of distribution Johnson noise or *white noise.*

   Figure 12

   > **Fig. 12.** "Johnson" or "white noise" melody. Pitches are randomly selected from a normal distribution centered on middle C (the most likely pitch).

   Since melodies are organized according to both pitch proximity and central tendency, melodies exhibit a sort of intermediate character between Brownian and Johnson fluctuations. Incidentally, Johnson noise has a so-called power distribution of $1/f^0$, whereas Brownian noise has a power distribution of $1/f^2$. When these two principles are combined, the resulting power distribution

approaches 1/f -- the so-called fractal distribution (Voss & Clarke, 1978; Gardner, 1978). Voss and Clarke (1975) have shown that melodies exhibit a power distribution similar to 1/f noise. While there are a number of natural phenomena that exhibit this distribution, there is nothing particularly magical about this observation.

3. **Ascending Leap Tendency/Descending Step Tendency**. In general, melodies tend to exhibit relatively rapid upward movements (ascending leaps) and relatively leisurely downward movements (descending steps). The reason for this asymmetry is not known. However, it is interesting to note that a similar phenomenon can be observed in the pitch of speaking voices. Researchers who study the "melody" of speech have observed that the initial part of an utterance tends to ascend rapidly, and then the pitch of the voice slowly drops as the utterance progresses. Linguists call this phenomenon *declination* and attribute it to the fall in sub-glottal air pressure as the lungs deflate (Pike, 1945; Lieberman, 1967; 't Hart, Collier & Cohen, 1990). Fig. 13 shows a randomly generated "melody" that is constrained only by an asymmetrical distribution favoring ascending leaps and descending steps. The melody behaves as a modified random walk, and so like Fig. 11 would inevitably drift out of range.

Figure 13

Fig. 13. Random melody based on asymmetrical distribution favoring descending steps and ascending leaps.

In an ideal world, these actual musical patterns would lead to the following subjective expectations:

1. **Pitch Proximity**. Listeners would expect an ensuing pitch to be near the current pitch.
2. **Regression-to-the-mean**. As the melody moves further away from the mean or median pitch, listeners would expect the next pitch to move closer to the mean.
3. **Downward Steps**. Listeners would expect most intervals to be descending steps.

Instead, experienced listeners show the following expectational tendencies:

1. **Pitch Proximity**. Listeners expect an ensuing pitch to be near the current pitch.
2. **Post-skip Reversal**. Experienced listeners expect a large interval to be followed by a change of direction.
3. **Step-Inertia**. Experienced listeners expect a small interval to be followed by a subsequent small interval in the same direction.

Like the pacific bull-frog, experienced listeners to Western music rely on patterns that are serviceable, but not exactly right.

Narmour's Theory of Melodic Organization

Note that these expectations conform very well to a theory of melodic organization proposed by Eugene Narmour (1990, 1992). Narmour proposed five predispositions that affect implicative melodic continuations (see Schellenberg, 1996 for a summary description). Two predispositions are

central to Narmour's implication-realization theory. The first is *registral direction* and the second is *intervallic difference.*

Studies by Cuddy and Lunney (1995) and Schellenberg (1996, 1997) have shown that Narmour's original theory can be simplified without loss of predictive power. Schellenberg (1997) in particular was able to show that Narmour's theory could be reduced to just two principles. One is the pitch proximity principle. The second principle is a combination of Narmour's registral direction and registral return dispositions. However, an analysis by von Hippel has shown that these phenomena can be accounted for by regression to the mean.

Similarly, earlier work by Rosner and Meyer (1982) and by Schmuckler (1989) had shown that listeners' responses are consistent with the notion of gap-fill. However, subsequent statistical analyses by von Hippel has established that the appearance of gap-fill is wholly attributed to regression-to-the-mean.

Narmour proposed that these expectations are somehow innate. At face value, the experimental research suggests that the expectations are learned, and that the expectation heuristics used by listeners are just approximations of structural properties present in the music itself.

Theoretically, it is possible that cause and effect might be reversed in the above account. It is possible that the organization of music has been shaped by *a priori* expectational tendencies rather than *vice versa*. That is, it is possible composers *intend* to create music conforming to *post-skip reversal,* but then somehow erroneously

construct melodies shaped by regression-to-the-mean instead.

This view is not very plausible, however. Regression-to-the-mean is a property of all distributions that exhibit a central tendency. The vast majority of distributions in nature show such central tendencies, so regression-to-the-mean is found wherever one cares to look. Moreover, there is a plausible explanation for why distributions of musical pitches would display a central tendency. When singing, vocalists find that it is physically easier to perform near the center of their range; both high and low notes are more difficult to sing. Similarly, most instruments are easier to play in some central register.

Scale Degree Expectations

Having examined pitches, interval sizes, and up/down contours, let us return again to consider the perception of scale degree. As we have seen, the key to understanding expectation begins by identifying patterns in the music itself.

In the first instance, we should consider the simple *event frequencies* for scale degrees. Like pitches, not all scale degrees occur with the same frequency. Bret Aarden (in preparation) has produced scale degree distributions based on a large sample of musical melodies. Figures 14 and 15 show the frequency of occurrence for works in major keys (first graph) and for minor keys (second graph). Both graphs are normalized by transposing all works so the tonic pitch is C.

Figure 14

Fig. 14. Distribution of scale tones for a large sample of melodies in major keys. All works were transposed so the tonic pitch is C; all pitches are enharmonic.

Figure 15

Fig. 15. Distribution of scale tones for a large sample of melodies in minor keys. All works were transposed so the tonic pitch is C; all pitches are enharmonic.

For both major and minor keys, the most common pitch is the fifth scale degree (dominant). In the major key, the second most common pitch is scale degree one (tonic) followed by scale degree three (mediant). In the minor key, the order of the tonic and mediant is reversed. Scale degrees four and two are next most common, followed by scale degrees six and seven. The non-scale or chromatic tones occurring least frequently.

The distributions shown in Figs. 14 and 15 are not merely an artifact of the aggregate of a large number of musical works. As it turns out, the scale degree distribution for most individual musical works are very similar to those shown in the figures. For example, the pitch-class distribution for J.S. Bach's Fugue No. 1 from the first book of the Well Tempered Clavier correlates with the aggregate major key distribution at +0.90. Such high correlations

turn out to be typical (Huron, 1992). Any musical passage written in a major key, that does not modulate to a different key for a prolonged period, will also show a strong positive correlation between its scale degree distribution and the aggregate distribution shown in Fig. 14. Similarly high correlations occur between works written in minor keys and the minor key distribution shown in Fig. 15 -- although the correlations tend to be lower for the minor keys compared with the major keys.

In an ingeneous set of experiments, Aarden (2002; in preparation) has shown that listeners' expectations conform to these distributions. Aarden established this by collecting reaction-time measures in a continuous listening task. Listeners were asked to press one of three keys (up, down, same) indicating the pitch-movement of successive pitches in various melodies. When listeners correctly anticipate an ensuing note, this is reflected in a faster reaction time. Conversely, when listeners are less certain of an ensuing note, this is reflected in a slower reaction time. When the data were collapsed according to scale degree, Aarden found that average reaction times were inversely proportional to frequency of occurrence. That is, listeners were faster when responding to scale degrees that occur more frequently in real music.

In a follow-up experiment (Aarden, in preparation), Aarden collected data only for the last note in a melody. Listeners heard 80 unfamiliar tonal folk melodies and watched a numerical counter count-down the number of notes remaining in the melody. When the final note appeared (count zero), listeners responded to the pitch contour (up/down/same) as quickly as possible. In this case, Aarden found a somewhat weaker correlation between the

average reaction times and the frequency of occurrence of various scale degrees. However, Aarden found a very high correlation between the average reaction times and the frequency of occurrence of final tones in a large sample of folk songs. That is, listeners were faster when responding to scale degrees that occur most frequently as the terminal pitches in a melody. Aarden's results imply that listeners maintain a different expectational "set" or "schema" for melody-final tones compared with ordinary melody tones.

Key Profiles

Aarden's work has provided an important clarification of a well-know experiment by Carol Krumhansl and Ed Kessler (Krumhansl & Kessler, 1982). Krumhansl and Kessler exposed listeners to a key-defining context, such as an ascending scale followed by a cadential harmonic progression. They then played an isolated "probe" tone, and asked listeners to rate how well the tone fits with the preceding context. They repeated this task using all twelve pitch classes and applied this procedure for both the major and minor key contexts. The results are shown in Figures 16a and 16b.

Figure 16

Fig. 16. Krumhansl and Kessler "key profile" for major context.

Krumhansl and Kessler "key profile" for minor context.

For a number of years, it was recognized that the Krumhansl and Kessler key profiles are similar (but not identical) to the frequency of occurrence for scale degrees in the respective major and minor key contexts. The principal difference is that the tonic is rated more highly in the Krumhansl and Kessler (K&K;) profiles. In addition, the second and fourth scale degrees (super-tonic and sub-mediant) are rated significantly less highly. In this, the K&K; distributions more closely resemble the distributions of pitch classes occurring at the ends of melodies, rather than the distribution of all pitch classes.

Aarden noted that since the probe-tone method stops the sequence of tones, listeners may tend to perceive this moment in terms of *closure*. In effect, rather than answering the question "how well does this tone fit with the preceding sequence of pitches?", listeners are answering the question "how well does this tone complete the preceding sequence of pitches?" (see also Butler, 19XX). Listeners' responses resemble the distribution of melody-final pitches much more than the general distribution of pitch classes. Using a multiple regression analysis, Aarden showed that the Krumhansl and Kessler key profiles can be fully accounted for by a combination of the general pitch-class distribution and the melody-terminating pitch-class distribution. More precisely, the distribution of melody-terminating pitch classes accounts for roughly 85 percent of the variance in the Krumhansl and Kessler key profiles, whereas the remaining variance (roughly 15 percent) is accounted for by the general pitch-class distribution.

Krumhansl has long argued that listeners are sensitive to the frequency of occurrence of various scale degrees, and

that learned mental schema arise for major and minor contexts (Krumhansl, 1990). However, discrepancies between the probe-tone profiles and the frequency distributions for actual works made Krumhansl's empirical evidence appear equivocal. Aarden's work brought clarity to the experimental data by showing that different schemata are employed for terminating pitches versus in-stream pitches, and that Krumhansl's experimental data are confounded by the perceptual closure that tends to accompany the probe-tone method. Once the distinction is made between in-stream and terminating pitch-class schemata, Aarden's work reinforces the view that listeners are indeed sensitive to the frequency of occurrence of pitch-classes.

Exposure Effect and the Pleasures of the Tonic

A favorite game musicians play involves performing a passage that provides a strong sense of key, and then to walk away from the music after playing the seventh scale degree or leading-tone. Most listeners find this experience grossly unsatisfying -- bordering on the intolerable. The music is left "hanging." By contrast, one can end on the tonic pitch and evoke a considerable sense of pleasure. What accounts for the psychological pleasure evoked by the tonic pitch?

In the first instance, not all tonic pitches evoke a sense of pleasure. When played as a passing tone in the context of a dominant harmony, the tonic will sound unstable and transient. The tonic pitch evokes the greatest pleasure

when it terminates a phrase or passage. That is, the pleasure of the tonic is linked to closure.

In the second instance, the tonic pitch is not alone in its capacity to evoke pleasure at moments of closure. The third (mediant) and fifth (dominant) scale degrees can also evoke a pleasant sense of closure -- although the pleasure evoked is often less than for the tonic pitch. In some jazz styles, ending on the sixth (sub-mediant) or even the second (super-tonic) scale degrees is often satisfactory.

As we have seen, the tonic is the most common way to end a musical passage. We might suppose that musicians choose to place the tonic at terminal moments because it sounds the most pleasant. But like all correlations, it is possible to confuse cause and effect. What if the pleasantness arises because the tonic is the most common terminal pitch?

Psychologists have documented, in innumerable ways, a tendency for people (and animals) to prefer the familiar (see review by Bornstein, 1989). Researchers have established that people have a preference for the "average" face. Similarly, Moreland and Zajonc (1977, 1979) carried out a set of experiments where subjects were exposed to various stimuli, such as complex polygons and Japanese ideographs. The stimuli were presented in such a way that the participants were unaware that some of the stimuli were being presented repeatedly. After an initial training period, the participants were exposed to another set of stimuli that contained both previous and novel stimuli. The subjects were asked to indicate whether they had seen the stimuli before, and were also asked which stimuli they preferred. A distracter task was included as part of the

experiment. Either due to the distracter and/or because of the complexity of the stimuli, the subjects were rather poor at discriminating between novel and familiar stimuli. However, in all experiments, subjects showed a marked preference for the more familiar stimuli. This preference for the familiar is referred to as the *exposure effect*.

In one of the Moreland and Zajonc experiments, tones of different frequencies were used. As in the case of the visual stimuli, listeners were unable to distinguish which frequencies they had been previously exposed to. Nevertheless, they showed a distinct preference for the most frequently occurring pitches. For musicians, this may not look like a very impressive result. Surely, the listeners were tending to assume that the most frequently heard pitch is the tonic. They preferred these tones *because* they heard them as tonics. That is, "tonality" would seem to explain the preference.

This interpretation is possible, although not perhaps very plausible. The phenomenon of preferring the most frequent stimulus is a general psychological phenomenon that has been observed with a wide variety of stimuli -- including both visual and auditory. Should we conclude that "tonality" is a fundamental phenomenon that operates in sequences of faces and polygons as well as tones? On the contrary, the experimental results suggest that the *exposure effect* is the more fundamental phenomenon. Listeners' preference for the tonic is more parsimoniously explained by appealing to the exposure effect rather than tonality.

Another reason for supposing that tonality is caused by the exposure effect, rather than vice versa, is that the effect is

not limited to isolated tones. Wilson (1975, 1979) carried out dichotic listening experiments in which various melodies were presented in one ear while a story was recited in the other ear. Subjects were required to follow the story line against a written text. The written distractor task was highly successful in getting listeners to ignore the melodies: in a subsequent recognition test, listeners performed at chance levels when asked to identify which melodies they had been exposed to. Nevertheless, listeners exhibited a preference for the melodies they had heard in the original exposure task. That is, entire melodies were preferred in a manner analogous to individual tones.

In a later discussion of expectation-evoked emotions (Part IV), an explanation will be offered for the origins of the exposure effect.

Expectation and Enculturation

The theory advocated in this study is that musical expectations arise from statistical learning through simple exposure to music. The results of Saffran *et al* (1996, 1999) provide strong evidence for statistical learning in tone sequences. But Saffran's experiments do not relate statistical learning to listeners' expectations. On the other hand, the work of von Hippel (2001) shows that the statistical properties of actual melodies are strongly correlated with the melodic expectations of listeners. But von Hippel's work does not demonstrate that the melodic expectations arise from statistical learning *per se*.

At the moment, there is unfortunately no direct experimental evidence testing the notion that listeners learn to infer statistical patterns from their past listening

experiences and use these statistical properties to form musical expectations. Nevertheless, the existing evidence is suggestive. In the absence of direct evidence, we can describe further experimental results that converge with this interpretation. Two pieces of converging evidence would be especially helpful. First, it would help to show that people from different cultural backgrounds exhibit different expectations when listening to the same music. Secondly, it would help to show that the expectations listeners exhibit reflect the statistical properties of the music found in their background cultures.

Consider first evidence that people from different cultural backgrounds exhibit different expectations when listening to the same music. In 1999, von Hippel, Huron and Harnish carried out an experiment that reveals how dissimilar expectations can be for different groups of listeners. The experiment contrasted the expectations of American musicians with Balinese musicians.

Both groups of musicians listened to a traditional Balinese melody played on a *peng ugal.* The Balinese musicians were highly familiar with the genre, whereas the American musicians indicated that they had little or no previous experience with traditional gamelan music. None of the participants had heard the test melody prior to the experiment. Each musician was tested individually using a betting paradigm.

The experimental apparatus consisted of a loudspeaker through which a sound recording of the melody could be heard, a digital keyboard sampler which reproduced the sound of the *peng ugal* and which was available to the musicians for consulting, a computer monitor that

displayed a limited set of notes from the melody using a numerical notation, and a physical mock-up of the instrument on which listeners could place bets using poker chips.

The goal of the task was for participants to place bets on each successive pitch of the melody and to attempt to acculumlate the greatest aggregate winnings. Bets placed on the correct pitch were rewarded ten-fold. Bets placed on the incorrect pitch were lost. Each participant was tested individually.

The participant heard the first note of the melody and the pitch was indicated on the computer monitor. The participant was then invited to bet on what they thought would be the likely second note. Having placed their bets, the actual second note would be reveallled, the winnings tabulated, and a sound recording of the melody played stopping before the third note. The participant was then invited to bet on what they thought would be the likely third note. This process was repeated until the entire 34-note melody was reveallled.

Throughout the experiment, participants could see the notation up to the current point in the melody, and could try out different continuations using the digital keyboard sampler.

In general, the results between the American and Balinese musicians were quite striking. Starting with a nominal grub-stake of $1.50, by the end of the melody, the most successful Balinese musician had amassed a fortune of several millions of dollars. The best American musicians failed to do as well as the worst Balinese musician.

Moreover, several American musicians went bankrupt during the game and had to be "advanced" a new grub-stake.

In the post-experiment interview, it was determined that all four Balinese musicians had been raised in religious homes where gambling was actively discouraged. So the differences between the American and Balinese participants cannot be ascribed to greater gambling experience for the Balinese participants.

Fig. 17 shows summary information for the American and Balinese listeners. When a person is uncertain of the outcome, they will tend to spread their bets over many more notes than if they are more confident of the likely outcome. A simple way to measure uncertainty is via entropy. Fig. 17 shows the average entropy for the American and Balinese listeners at each point as the melody unfolds. A rough approximation of the melodic pitches is provided using Western notation.

Figure 17

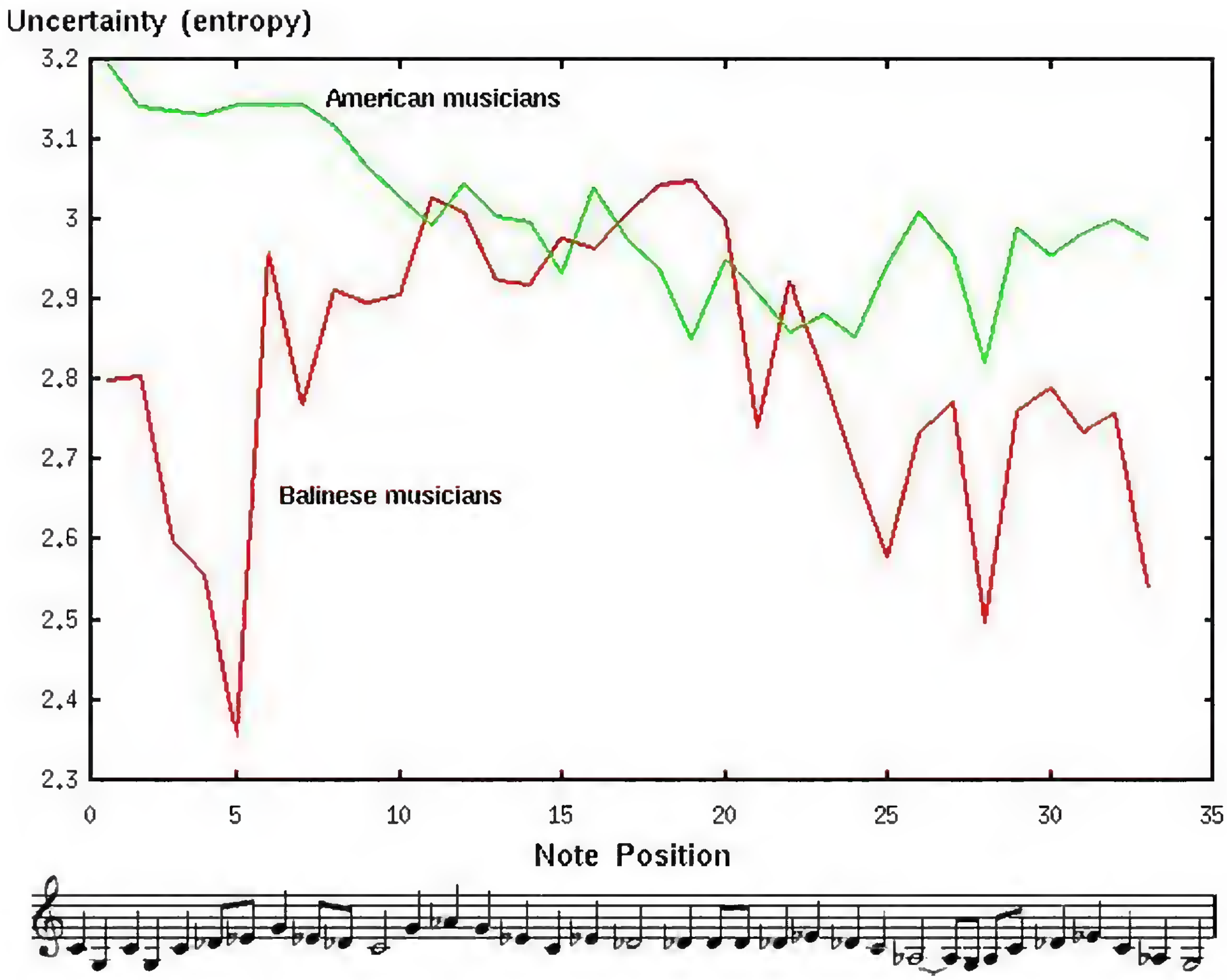


Fig. 17. Average moment-to-moment uncertainty for Balinese and American musicians listening to an unfamiliar traditional Balinese melody. Uncertainty is plotted as entropy, measured in bits. In general, Balinese listeners show less average uncertainty. Note positions correspond with underlying notational rendering. N.B. Notation shows only approximate pitch levels.

The graph shows that, on average, the Balinese listeners were nearly always less uncertain of possible future continuations than the American listeners. Since the American and Balinese musicians were matched for age, sex, and general musical experience, these differences are likely to have arisen due to the Balinese musicians greater familiarity with traditional Balinese music.

Consider next the issue of demonstrating that listeners' expectations reflect the statistical properties of the music found in their background cultures. Perhaps the best evidence in support of this can be found in research pertaining to structural tonality. Among Western listeners, a wealth of experimental data shows that the simple frequency of occurrence of various pitch classes plays a significant role in tonality perception (Cuddy, 1993; Cuddy & Baderstscher, 1987; Lamont, 1998; Oram & Cuddy, 1995). More importantly, evidence consistent with structural tonality has been observed in non-Western musical practices -- such as in classical Indian music (Castellano, Bharucha & Krumhansl, 1984), in Balinese music (Kessler, Hansen & Shepard, 1984), and in Korean *p'iri* music (Nam, 1998). In the case of traditional Korean music, for example, the most frequently occurring pitch also tends to terminate breath-delimited phrases, and also coincides with the pitch identified by Korean musicians as the central pitch of the scale. Moreoever, these pitches change systematically with respect to different Korean modes, and even with transposed pitch sets.

In the Castellano, Bharucha and Krumhansl study (1984), both American college students and Indian listeners were exposed to samples of North Indian music, and tested using the probe tone technique. Both groups responded in ways that echoed the frequency of occurrence of pitch classes in the musical samples. However, in carrying out a multiple regression analysis, Castellano *et al* were able to remove the variance associated with the exposure frequencies and examine the residual variance. It was found that the residuals for the Indian listeners correlated with a hierarchy of pitches in established Indian music

theory (Jairazbhoy, 1971), whereas the residuals of the American listeners showed no such correlation. In effect, Castellano *et al* demonstrated that (1) Indian listeners were responding in a way that combined long-term statistical features engendered through years of listening to Indian music, plus short-term statistical properties related to the actual exposure sample, whereas (2) American listeners responded in a way that was consistent with the statistical properties of the exposure sample. Note, however, that the American listeners were showing clear evidence of statistical learning for the Indian musical excerpts.

In both the Castellano, Bharucha and Krumhansl and the Kessler, Hansen and Shepard studies, listeners were instructed to respond according to the "stability" or "goodness of fit" of the probe-tone. Although there is ample evidence consistent with statistical learning, one cannot claim that listeners responses represent *expectations* only, and so the evidence for statistically learned expectations remains indirect.

Taken together, all of these studies lend support to the view that musical expectations arise from statistical learning through (both short-term and long-term) exposure to music.

Cadences

Not all musical moments are equally predictable. Possibly the most cliché aspect of music can be seen in how things end. Theorists have long recognized that cadence points tend to be organized in a stereotypic fashion. The stereotypes of musical closure can be readily observed in

Landini cadences, in dominant-tonic harmonies, and in innumerable pre-cadential formulas, such as suspensions, the use of augmented sixth chords, and pre-cadential second inversion chords (see, e.g., Kramer, 1982). Figure 17 shows that Balinese listeners tend to become less uncertain of the next note as the end of the melody is approached.

Consider, for example, a sample of 300 German folksongs from the Essen Folksong collection (Schaffrath, 1995). A simple calculation might examine the information content (in bits) of pairs of successive scale degrees. For example, among the highest probability events is the dominant pitch followed by a repetition of the dominant (4.1 bits). By contrast, a low probability (high information) sequence consists of the lowered seventh followed by the raised seventh (13.4 bits). Over the complete sample of 300 folksongs, the average information content for scale-degree successions is 5.52 bits (S.D. of 1.42). However, the information content for the final two notes of each phrase is 5.08 bits (S.D. of 1.29). Such patterns are ubiquitous throughout music and can be observed in such disparate repertoires as Gregorian chant, Pawnee music, and Bach chorale melodies (Manzara, Witten & James, 1992).

Apart from closure and cadences, a number of organizational patterns are evident in many musical genres. Gjerdingen (1988), for example, has identified a number of widespread clichés associated with the classical style.

Expectation in Time

So far we have been considering only pitch-related expectations. Listeners not only form expectations about what future events may occur, but also *when* they occur. Caroline Palmer and Carol Krumhansl carried out a set of probe-tone studies to determine when listeners most expect events to happen. Palmer and Krumhansl (1990) presented stimuli that created particular metric frameworks, like 4/4 and 3/4. Following a meter-defining sequence, there was a pause, followed by a tone. Listeners were asked to judge the "goodness of fit" for each tone. Listeners assigned the highest values to those tones whose onsets coincided with the most important beats in the metric hierarchy, followed by the lesser beats, followed by the half-beat divisions, followed by tones that did not coincide with any beat.

Mari Riess Jones has proposed that the metric hierarchy can be understood as a structure for rhythmic attending. Auditory attention is directed at moments in time. That is, when listening, auditors do not pay attention equally at all moments. In *rhythmic attending*, Jones notes that the listener's attention is most acute at strong metric positions. That is, the metric hierarchy corresponds to a sort of temporal expectation framework.

Consider the following experiment carried out by Jones, Moynihan, MacKenzie and Puente (in press). Listeners heard an initial tone, followed by 12 "distractor" tones, followed by a comparison tone. The task of the experiment was for listeners to judge whether the comparison tone was higher or lower in pitch than the initial tone. In the following example, the first pitch (half-note B) is the initial tone, and the final pitch (half-note A#) is the comparison

tone. The intervening tones are random distractor tones that increase the difficulty of the task.

Figure 18

Fig. 18. Typical stimulus used in Jones, Moynihan, MacKenzie & Puente (in press). Listeners heard a standard tone, followed by twelve interference tones, followed by a comparison tone. Listeners were asked to judge whether the comparison tone is higher or lower than the standard tone. The temporal position of the comparison tone was varied so that it would occur earlier or later than expected. See also Fig. 19.

Jones *et al* manipulated the precise temporal position of the final comparison tone. In some trials, the onset of the tone coincided with the precise downbeat (position 3). Other trials were slightly ahead (position 2) or slightly delayed (position 4) compared to the downbeat. Yet other trials were considerable ahead (position 1) or delayed (position 5) compared to the downbeat. Jones *et al* found that the accuracy of pitch-comparison judgments depended on the precise temporal placement of the comparison tone. Listeners were most accurate in their judgments when the comparison tone coincided with the presumed downbeat. As the tone deviated from this position, perceptual judgments were degraded:

Figure 19

Fig. 19. Effect of temporal position on accuracy of pitch judgment. (See also Fig. 18.) Jones, Moynihan,

MacKenzie & Puente (in press) showed that pitch judgments are most accurate when the tone judged occurs in an expected temporal position (position 3).

This research reinforces and extends the general principles we have already seen operating with regard to auditory expectation. Specifically,

1. **Expectations facilitate perception.**

   It is not simply the case that expectations prepare an organism to take appropriate action. In the case of temporal expectations, we see that listeners expect to receive information at certain times. The listener may not know *what* is going to happen, but might nevertheless anticipate the moment when the information arrives.

   One can imagine a number of ways in which accurate expectations facilitate perception. The prospect of perceiving something with greater accuracy could well be responsible for encouraging an organism to attempt to form accurate expectations about the future. In this sense, temporal expectations are akin to the *orienting response* -- a behavior that improves perception.

   In addition, expectations can be viewed as preparations for appropriate motor behaviors.

2. **Expectations are shaped by context.**

   As in the case of pitch perception, rhythmic expectations are related to the context. Some contexts are quite general, as when we experience music in simple-duple meter, or compound-triple meter. At the

other extreme, we may expect a particular temporal organization because of extensive familiarity with a particular rhythm or musical work. That is, rhythmic expectations may arise through veridical contexts.

It is also possible that listeners form schematic expectations that are culture- or genre-related. Consider, for example, the *siciliano* -- a leisurely baroque dance form. The *siciliano* is generally in 6/8 meter, although occasionally it is found in 12/8. In addition to this compound-duple metric framework, there are stereotypic rhythms the occur in this form and that contribute to the stylistic cliché for the *siciliano*. The most distinctive feature is the dotted-eighth/sixteenth figure that begins the measure, and the quarter-note in the mid-measure position, followed by either an eighth-note or two sixteenths:

Figure 20

Fig. 20. Two rhythmic patterns commonly found in *siciliano* dance forms.

Schubert's famous Christmas carol, *Stille Nacht* ("Silent Night"), exhibits the distinctive *sciliano* rhythm. Below is a cumulative onset histogram for a sample of bars from various *siciliana,* showing the relative frequency of occurrence for various points in the 6/8 metric hierarchy.

Figure 21

Fig. 21. Cumulative onset histogram for a sample of bars from various *siciliana* movements, showing the relative frequency of occurrence for various points in the 6/8 metric hierarchy.

Once established, listeners readily expect the rhythm. In this case we can see that it is not simply the strict hierarchical metrical frameworks that influence a listener's temporal expectations. In addition to these *metric* expectations, listeners can also form distinctly *rhythmic* expectations which can employ non-regular duration patterns. Expectations can be tailored for different rhythms: *sambas*, *tangos*, rock *back-beats*, and so on. Similarly, complex African rhythms can evoke specific temporal expectations for those listeners who are familiar with them. [3]

3. **Temporal expectations are learned.**

Although no one has provided a formal demonstration, it is quite likely that rhythmic expectations are shaped by the same statistical learning of the auditory environment that we've seen for pitch. The reason why periodic pulse and meter are common in music is that these patterns are the easiest patterns for which brains are able to form expectations. In this regard, the metric hierarchy is truly analogous to a scale or scale hierarchy. Metric positions provide convenient "bins" for expected stimuli.

While periodicity is helpful for listeners, periodicity is not necessary in order to form temporal expectations. It is important only that the listener be experienced with the temporal structure, and that some element of the

temporal pattern be predictable. An illustration of this point can be found in the expectation for "bouncing" rhythms (see Fig. 22). Although the sound of something bouncing is not periodic, the inter-bounce interval shortens predictably as the bouncing continues and so listeners are able to predict, to some degree, the temporal sequence of events. In music, this accelerating rhythm can be found in Tibetan monastic music (where it is frequently played on cymbals). In Western music, there is no known instance of this accelerating rhythm prior to the twentieth century.

Figure 22

Fig. 22. Schematic representation of accelerating onsets characteristic of the sound produced by a bouncing object. Although the pattern is not metrically regular, it is nevertheless predictable.

Long-Range Contingent Expectations

To this point, our discussion of contingent expectations has focussed on comparatively short-range phenomena. Typically, we have been considering the repercussions of some event only on the immediately ensuing event. However, it is often the case that an event will have a greater impact on somewhat distance events than on neighboring events. Mari Riess Jones has assembled a wealth of data illustrating the hierarchical nature of

auditory attending in time. Expectations in time appear to exhibit a range of local to global effects (see Jones, 1992).

Earlier we saw how information theory can be used to characterize short-term conditional probabilities. A branch of information theory known as "m-dependency" theory provides useful ways to characterize long-term statistical relationships between events (see Wong & Ghahraman, 1975). In English text, we know that the letter "q" tends to constrain subsequent letters -- increasing the likelihood of an ensuing letter "u". But can a letter not influence the occurrence of letters that follow at a further distance?

Figure 23 shows the interdependence of successive characters in English text. The X-axis indicates the number of characters following a given target character. The Y-axis measures the dependency (in bits). As can be seen, the strongest effect is evident for a single character. This captures, for example, the strong influence the letter "q" exerts on the ensuing character. As the distance increases, the influence decreases exponentially. The lower line in Figure 24 shows the dependencies for randomly scrambled English text. The only influence that a randomly rearranged character can have on the ensuing character relates to the overall frequency of occurrence for various letters. This line establishes a random base-line that is useful for comparison purposes. The figure shows that the future influence of an individual letter in English text declines to zero at a distance of about 6 letters.

Figure 23

Fig. 23. Graph showing the influence in English text of one letter on the presence of another letter displaced by *n* characters. Consecutive letters (*n=1*) have considerable dependency. At a distance of about 6 letters the presence of a given letter has little measureable influence on a later letter. Independence is measured as entropy (in bits). From Simpson (1996).

Working at the University of Waterloo, Jasba Simpson applied m-dependency theory to the analysis of note-dependency in music. Simpson examined four musical works: The works are Debussy's *Syrinx* for solo flute, Bartók's *Unison* for piano, Bach's *Prelude I* in C major from the first volume of the Well-Tempered Clavier, and Bach's *Allemande* from one of the six flute sonatas. The results of the analyses are shown in Fig. 24. Once again, the graphs plot the distance over which one note influences another note.

Figure 24

Fig. 24. Interdependence graphs for four musical works. Claude Debussy's *Syrinx* for flute. Bela Bartók's *Unison* for piano (from *Mikrocosmos*), Johann Sebastian Bach's *Prelude I* in C major from Volume 1 of the Well-Tempered Clavier, and Bach's *Allemande* from one of the sxi flute sonatas. The graphs show long term note dependencies. From Simpson (1996).

Both the Debussy and Bartók works exhibit the exponential decay typically found when the dependencies are relatively

short-range. The strongest contingencies are evident when the events are close. As the notes grow further apart they exhibit less of a statistical influence on one another. In the case of the two Bach works, however, there are significant peaks evident at the higher probability orders. Note especially the graph for the Bach C major Prelude. The dependencies between successive neighbors is relatively small. Instead, the greatest influence is apparent at 8 and 16 note separations. The reason for this relationship is obvious when looking at the score (see Fig. 25).

Figure 25

Fig. 25. Opening measures from Johann Sebastian Bach's *Prelude I* in C major from Volume 1 of the Well-Tempered Clavier. Repetitive patterns are evident at 8 and 16 notes distance. These dependencies can be seen in the corresponding graph in Figure 24.

Throughout this piece, Bach establishes series of parallel compound melodic lines. The two voices in the bass staff are notated clearly enough, but even the seemingly singular series of sixteenth notes in the treble staff is perhaps better regarded as three independent voices. Clearly, each pitch has a strong relationship to pitches 8 and 16 notes distant. For example, the highest pitch E5 in measure 1 is connected perceptually to the pitch F5 in the second measure (Bregman, 1990; Schenker, 1906).

This sort of organization is relatively less common in the case of language -- although not entirely absent. For example, such long range dependencies can be observed in

poetry with regular rhyme schemes. The statistical methods provided by m-dependency theory allow us to measure and characterize such relationships.

The fact that musical works exhibit long-term dependencies raises two questions. First, do listeners form corresponding expectancies where the implicative events are some distance removed from the expected consequence? Second, since the long-range patterns identified above are associated with individual works, do listeners quickly form new expectancies that are tailored to the unfoleding events of a musical work?

Relatively little experimental research has addressed either of these questions. Richard Aslin at the University of Rochester has carried out a series of studies where sounds are contingent on subsequent sounds, but the two sounds are separated by a statistically unrelated sound. Alsin *et al* have studied successions of synthesized vowels, consonants, and pitched tones. The results of these experiments are complicated. For some kinds of stimuli, listeners form appropriate expectations, whereas listeners fail to form useful expectations for other kinds of stimuli. Moreover, Aslin and his colleagues have also performed the same experiments with cotton-top tamarins and shown that these primates exhibit a different pattern in forming suitable expectations.

It is not simply the case that tamarins are unable to form some expectations that humans readily do. For some stimulus patterns, tamarins succeed in forming appropriate expectations where human listeners fail. These inter-species differences are tantilizing, and might ultimately

prove to be linked to special speech-related mechanisms for processing sound sequences.

In any event, the research pertaining to long-range expectations appears to be consistent with past experimental results -- suggesting that listeners form expectations that only approximate the true underlying patterns of contingent probabilities.

Quick Study

The second question posed above asks whether listeners rapidly form expectations that are uniquely tailored to the unfolding events of an individual musical work. The above results suggest that listeners adapt their expectations to individual musical works. As the events of the piece unfold, the work itself engenders expectations that influence how the remainder of the work is experienced. This view was proposed by Meyer (1956). As we saw earlier, Castellano, Bharucha and Krumhansl (1984) have provided experimental evidence that listeners do indeed adapt relatively rapidly to music not previously encountered.

This phenomenon of rapid adaptation was anticipated in early research in information theory. Most notably, Coons and Kraehenbuehl proposed an adaptive probability model for experiencing music as it unfolds as early as 1958. (Coons & Kraehenbuehl, 1958; Kraehenbuehl & Coons, 1959). Kraehenbuehl and Coons imagained that a listener's statistically-shaped expectations would become better adapted to a musical work as the amount of exposure increased. A listener would begin the listening experience with expectations reflecting broad or generalized probabilities arising from a life-time of musical exposure.

But as the musical piece progresses, the listener would build expectations that are engendered by events in the work itself. The ability to model such adapative probabilities was beyond the technology available in the 1960s. By the time the technology made such modelling feasible, music theorists had lost interest in information theory. No one has yet pursued such an adaptive modelling approach.

Schematic and Veridical Expectations

With repeated exposure, a listener can become highly familiar with a given musical work. In many instances, an entire musical work is committed to memory. Clearly, a listener has nearly "perfect" expectations for highly familiar pieces, such as *Happy Birthday*. At any given point in the work, the listener knows precisely what will happen next. Such seemingly "perfect" knowledge implies that no variability in expectation would be possible. At all points, the listener has complete knowledge of the ensuing events. When a work is perfectly known to some listener, what does it mean to have expectations? How does extreme familiarity with a single piece change the experience of listening to that piece?

Of course this knowledge is not entirely perfect. It typically requires several notes at the beginning of a work for the listener to gain confidence that the work is what they think it is. With just the first note, some element of doubt will exist. In addition, music typically contains repeated sections, and at particular structural points, the listener may be in doubt about the precise continuation. One piece of evidence in support of this claim can be found in the

sorts of memory errors often seen when amateur musicians play recitals or auditions. A nervous performer sometimes lapses into a memory "loop" where they play the same passage verbatim without taking a "second ending" or otherwise continuing as they should with the rest of the piece. In short, there can still exist points of uncertainty, even in highly familiar works.

A more compelling problem is how an experienced listener might continue to hear elements of uncertainty that are similar to those for listeners hearing the music for the first time. This paradox is sometimes referred to as *Wittgenstein's Puzzle* (see Dowling & Harwood, 1986; p.220). A classic example of this problem arises in the perception of the *deceptive cadence*. How, we might ask, can a deceptive cadence continue to sound "deceptive" when familiarity with a work makes the progression entirely inevitable?

One possible answer lies in an apparent bifurcation of the neurophysiological paths related to expectation. One path represents a low-level path where highly practiced patterns of exposure are coded. A second path represents a higher-level, less practiced pattern of exposure. In cognitive terms, these two different paths might correspond to the distinction between *schematic memory* and *veridical memory*. Veridical memory is memory for specific events, whereas schematic memory is memory for general patterns. The difference can be illustrated using two well-known English phrases:

> *Once upon a time ...*
>
> *Four score and seven years ago ...*

In the first example, the phrase "Once upon a time" can be found at the beginning of a large number of legends and fairy tales. Several continuations are possible:

> *Once upon a time there was a little girl named Little Red Riding Hood ...*
> *Once upon a time there were three bears ...*

The second example, "Four-score and seven years ago" is unique to Lincoln's Gettysburg address. There is only one expected continuation:

> *Four score and seven years ago, our fathers brought forth upon this continent a new nation ...*

Jamshed Bharucha has drawn attention to the applicability of these concepts to understanding musical expectation. Bharucha and his colleagues (1999) have shown that schematic-engendered responses are still evident in veridical listening tasks. For example, a deceptive cadence can still evoke a physiological response characteristic of surprise, even when the listener is certain of its occurrence. In effect, the fast (schematic) brain is surprised by the "deception" while the slow (veridical) brain is not.

The reason why schemas exist is to allow the brain to respond more quickly to particular situations. These schemas therefore reflect the most commonly encountered contigent expectations. That is, the schemas represent broadly enculturated aspects of auditory organization.

Of course, if a culture existed where nearly all dominant chords are followed by a submediant chord, then the *V-vi*

chord progression would no longer be perceived as deceptive. As long as the majority of dominant chords in a culture are not followed by the submediant, this progression will still retain an element of surprise.

By way of summary, in the above discussion we have distinguished three different levels or frameworks for expectations. *Schematic expectations* represent broadly enculturated patterns of events. *Veridical expectations* represent long-term patterns arising from repeated exposure to a single complex episode. *Adaptive expectations* represent dynamically up-dated patterns that quickly arise in the context of a novel exposure, such as the first-hearing of a musical work.

Origin of Schematic and Veridical Memory

A helpful question is to ask why the brain distinguishes between schematic and veridical information. Why are some things remembered or coded as general principles, while other things are remembered or coded as specific events?

This question can be rephrased in terms of so-called *episodic* and *semantic* memory. In general, it is more efficient to recall general principles rather than specific events. For example, it is simpler to remember that "Eric is untrustworthy" than to remember a series of past events that all seem to testify to Eric's untrustworthiness. When we are tempted to ask Eric to attend to an important task, it is faster and more efficient to access the general principle rather than ponder all of our past interactions.

When a person concludes that "Eric is untrustworthy", based on past experience, they are making an inductive inference -- forming a general proposition based on a finite series of observations. However, as we noted earlier, induction is itself fallible. In fact, we have seen instances where observations lead to the wrong inference. It is quite possible that Eric is indeed trustworthy. When he failed to show up as promised, he might have had to take his mother to the hospital and then did not have an opportunity to explain. When Pat relayed negative gossip about Eric, perhaps Pat was attempting to unjustly tarnish Eric's reputation so that Pat would be promoted rather than Eric.

Cosmides and Tooby (2000) have argued that retaining episodic memory is functionally essential. In effect, episodic memory allows us to revisit "the original data" in order to evaluate alternative hypotheses. If we simply retained the generalized semantic or schematic information ("Eric is untrustworthy") and discarded the original episodic or veridical information, then we would be unable to reconsider a possibly questionable inductive inference.

Clearly, the brain's ability to form generalizations is important. But it is also clear that the brain needs to retain some of the original observational data so that the credence of particular generalizations can be questioned, revised, or reinforced. Evolution has addressed the problem of induction by creating two parallel memory systems.

In the case of the auditory system, these systems are evident in listening schemas that represent current generalizations about the world of sound, as well as a

learned veridical system. Either system can be surprised. As we saw in the deceptive cadence, the schematic system is surprised while the veridical system is not. But it is also possible to arrange the reverse. In Fig. 26 a chimeric melody is shown that begins with the notes of "Three Blind Mice". However, at the end of the second measure, the continuation is inconsistent with "Three Blind Mice". The melody elides into "Mary Had a Little Lamb." The switch is surprising from a veridical perspective. But the pitch sequences themselves are commonplace, and so there is no schematic surprise.

Figure 26

Fig. 26. Example of a chimeric melody where one melody elides into another. At the end of the second measure, an experienced listener will experience a "veridical surprise". However, the pitch sequences themselves are commonplace, and so there is no schematic surprise.

Recovering from Wrong Notes in Improvisation

Another example of the relationship between veridical and

In music improvisation, the performer must be able to contend with unintended "accidents" -- slips that would normally be considered errors. Whether one is improvising a jazz chart or realizing a figured bass accompaniment, experienced musicians have been uniform in offering novice improvisors the advice of returning to the "wrong"

note and playing the passage again including the wrong note. The goal is to convince the listener that the note was not an error, but was intentional.

First, what do we mean by an improvised note being "wrong"? From an expectational standpoint the answer is straightforward: the note has a low probability of occurrence. Given its low likelihood, the initial appearance of the wrong note will inevitably sound jarring to the listener. However, repeating the passage will allow the listener to accommodate the errant note within a newly formed expectation.

In effect, the experienced improvisor establishes the "wrong note" as a normal part of a veridical passage. The performer can do nothing about the violation of the schematic expectation. In particular, the performer can do nothing to erase the original surprise evoked by the first appearance of the wrong note. However, by encorporating the passage as part of the work, listeners can be dissuaded away from the conviction that the performer has made a mistake.

Violations of schematic expectation are commonplace in music. However, violations of veridical expectations tell listeners that something is wrong -- that the performer has messed up. The performer has mis-played "the piece."

Anchoring and Tendency Tones

As we saw earlier, different scale tones are perceived to have different degrees of stability, with the most frequently occurring tones generally having the greatest stability. Intuitively, we tend to think of the less stable tones as

exhibiting some sort of tendency. For example, the leading-tone has a tendency to be followed by the tonic pitch.

Figure 27 was produced by Bret Aarden from the Ohio State University. Aarden simply measured the probability that certain scale tones would be followed by other scale tones. Some scale tones are highly constrained by what happens next. For example, the raised dominant is nearly always followed by the submediant pitch. Other tones, like the dominant, can be followed by a much greater variety of continuations. Figure 27 plots the information content (in bits) for each scale degree. If listeners acquire some knowledge of the probabilities associated with scale-degree successions, then this graph should correspond to our expectations of *tendency*. That is, those scale tones toward the right side of the figure will evoke a greater sense of "leading" or "tending".

Figure 27

Fig. 27. Scale tones for C major ordered according to the range of possible ensuing tones. "Flexibility" is measured as entropy (in bits). The dominant pitch (G) can be followed by many different pitches. By contrast, the raised dominant (G#) tends to severely constrain possible pitch continuations. (Calculated by Bret Aarden, 2001).

What Figure 27 doesn't show is that the strongest tendency tones lead to nearby tones. That is, typically, a tendency tone will cleave to a more stable tone that is just above or just below within the scale. It almost seems that the closer a less stable pitch is to a more stable pitch, the greater the

tendency for the less stable pitch to be followed by the more stable pitch. Recall that by "stable" here, we simply mean tones that have been learned to appear more frequently, that are preferred (due to the exposure effect) and that evoke less stress.

The importance of tendency tones, and the tendency to hear them as linked to more stable neighbors was vividly described by the University of Pennsylvania theorist, Leonard Meyer (1956; p.56). Meyer noted, for example, that "In the music of China non-structural tones take the name of the structural tone to which they move together with the word *pièn*, meaning "on the way to" or "becoming." That is, the tendency tones are named in reference to the "resolving" tone.

Anchoring and Embellishment

As we have seen, there is a strong tendency to perceive stimuli in terms of pre-existing formulas or schemas. However, this does not mean that we perceive only what we expect. We can often tell when a performer plays a wrong note; we can be surprised in music -- and we can be disappointed as well.

There is a weaker sense in which perceptions are assimilated into schemas. We may perceive that an event is not quite right, but still interpret this discrepancy in terms of a useful pattern -- such as a schema or prototype. Consider an example studied by Eleanor Rosch (1975). Rosch found that a line tilted 10° to the horizontal is perceived to be similar to a horizontal line. She also found that people judge a slightly tilted line to be more similar to

a horizontal line than the horizontal line is judged similar to the tilted line. In short, the horizontal line acts as a prototype that provides a cognitive reference point for the tilted line. The tilted line is perceived as a slight variant of the prototypic horizontal line. The tendency to interpret a stimulus as a variant of a prototype is called *anchoring.*

Krumhansl (1990) showed that, in a given key context (such as C major or G minor), the most stable tone is the tonic, followed by the other tones of the tonic triad, followed by the remaining diatonic scale tones, followed by the chromatic tones. In the perception of melodies, Bharucha (1984) demonstrated how less stable tones tend to become anchored to ensuing, more stable tones, that are close in pitch. For example, in the key of C major, the pitch D has a tendency to be anchored to either the neighboring C or E. Similarly the pitch D# has a tendency to be anchored to the nearest more stable pitch E. Bhuarcha asked listeners to judge whether two five-note melodic framents were identical. A single wrong note was introduced in many of the trials. Two sample trials are illustrated Fig. 28. In both trials, the target five-note melodic fragment consists of the pitches E4, G4, C5, D5, E5. Comparison passages "a" and "b" both introduce a single wrong note (B4 and F4, respectively). Listeners were much more likely to judge fragment "a" is identical to the target than passage "b". Bharucha argued that B4 tends to be better anchored to the ensuing pitch C5, and so it becomes less noticeable as a wrong note. By contrast, the F4 is not anchored to the ensuing pitch and so is more noticeable.

Figure 28.

Fig. 28. Experimental stimuli used in Bharucha (1984). Listeners were asked to identify whether the first and second five-note patterns were the same of different. The target passage (E-G-C-D-E) is the same. Comparison passage "a" was more likely to be mistaken from the target passage than comparison passage "b". Bharucha argued that the reason for the greater similarity is that the wrong pitch (B4) in "a" is anchored to the more stable subsequent pitch (C4), whereas the wrong pitch (F4) in "b" fails to be anchored to the ensuing pitch and so is more noticeable.

Conscious Expectations

To this point we have considered only with those aspects of expectation that are pre-verbal or unconscious in origin. It is also possible for listeners to develop conscious strategies arising from verbalizable knowledge. An example of such conscious expectations can be seen in the knowledge of sonata-allegro form. Sonata-allegro structure provides an organizational framework that knowledgeable listeners can employ in forming future expectations. An aware listener can use form-related sign-posts to orient herself or himself. For example, one might turn on the radio and hear a classical work already in progress. One might hear a plausible "first theme" followed by a plausible "second theme." By noting that no modulation occurred between the two themes, the knowledgeable listener could infer that

the performance is in the midst of the recapitulation section, and so the ending can be expected shortly.

Some music theorists have presumed that these kinds of large-scale form-related expectations are also present at an unconscious level. However, research by Vladimir Konecni has raised doubts about this assumption. Working in the Psychology Department at the University of California, San Diego, Konecni and his colleagues have shown that listeners are surprizingly insensitive to reorderings of musical segments (e.g., Gotlief & Konecni, 1985; Karno & Konecni, 1992). The original versions of musical works consistently fail to elicit a greater preference than altered versions for both musician and non-musician listeners. Similar results have been found by Nicholas Cook in the Music Department at the University of Southampton (Cook, 1987).

Once listeners become familiar with style-related clichés, it becomes possible to thwart or otherwise manipulate the normal expectations. A good example of this with respect to closure in Western art music can be found in Haydn's so-called *Joke Quartet*.

III. MUSICAL GENRES AND ENVIRONMENTAL CONTEXTS

In forming expectations about the world, it is easy for past experiences to become over-generalized. We may not realize that our expectations have value only in a specific narrow realm. When the context is wrong, otherwise useful information may prove false, misleading, or even harmful. As Cosmides and Tooby (2000) have noted, there are good

reasons why, in the evolution of cognitive processes, special mechanisms would be needed in order to limit the scope of learned information.

When listening to music, our expectations can change dramatically depending on the style or genre of the music. In *reggae*, for example, there is a strong likelihood that a dominant chord will be followed by a subdominant chord. But in Western classical music, this dominant-subdominant progression is much less common. If the experienced listener is to correctly anticipate the unfolding of acoustic events, then the listener must somehow bracket two different sets of expectations. By forming two different schemas, the listener is presumed to be able to hear the dominant-subdominant progression in one context (*reggae*) as a commonplace event, and in another context (*classical*) hear the same chord progression as somewhat surprising.

Music is not unique here. Social psychology provides innumerable illustrations of the effect of context on expectation. Norms of behavior are linked to particular social roles. For example, we comply with the family doctor who asks to take a look in our ears, but we would be dumbfounded if the same request were made by a sales clerk. As sociologists have noted, the wearing of distinctive uniforms is an important way of providing role-relate cues. These overt cues help us switch between different expectational sets or schemas.

We already have good evidence for the existence of different musical schemas. Perhaps the best documented difference is the distinction between major and minor modes (Krumhansl, 1990). Western listeners exhibit dramatically different expectations depending on whether

the music is perceived to be in a major or minor key. A single musical work may contain passages that switch between the major and minor modes. The existence of such works suggests that listeners are competent in switching schemas as the music unfolds. A further lesson arising from the major/minor distinction is that musically pertinent schemas are not simply restricted to different styles, genres, or cultures.

If our musical expectations change according to context, then a number of important questions arise: How many different musical schemas can a listener maintain? How fast are listeners able to identify the context and invoke the appropriate schema? When the context changes, how fast are listeners able to switch from one schema to another? What cues signal the listener to switch schemas? How do listeners learn to distinguish different contexts? How are the expectations for one schema protected from novel information that pertains to a different schema? How does a listener assemble a totally new schema? What happens when the events of the world straddle two different schemas?

Schema Selection

We might begin by asking how listeners know what schema to start with. We already know that an isolated tone tends to be heard by listeners as the tonic. But is this the tonic of a major or minor key? Following exposure to an isolated 2-second tone, listeners are more than three times as likely to expect a tone whose pitch is a major third above as a minor third above. This implies that Western listeners have a tendency to start by assuming a major mode. [4] It is

conceivable that a musically-pertinent schema may be invoked prior to the onset of any sound.

Once the music has begun, how fast are listeners able to recognize the musical context? In the case of music, dramatic changes in listeners' expectations arise depending on the style or genre of the music. Perrott and Gjerdingen (1999) have observed that listeners are very quick to identify different styles. When scanning the radio dial, listeners make split-second decisions regarding the style of music being played on each station. Perrott and Gjerdingen tested this observation by selecting random musical segments from samples of 10 different styles of music, including jazz, rock, blues, country & western, classical, etc. They showed that listeners are adept at classifying the type of music in just 250 milliseconds. With just one second of exposure, ordinary listeners' abilities to recognize broad stylistic categories is nearly at ceiling; that is, further exposure to the musical work does not lead to a significant improvement in style identification. If we assume that identifying a schema is tantamount to activating the schema, then these observations suggest that experienced listeners can activate a schema appropriate to the genre of music they are hearing in a very short period of time.

What about the phenomenon of *schema switching*? How rapidly can a listener switch from one schema to another? Although little research has been carried out pertaining to this question, suggestive evidence has come from the work of Krumhansl and Kessler (1982). Krumhansl and Kessler traced the speed with which a new key was established in modulating chord sequences. Modulations to related keys were "firmly established" within three chords lasting a few

seconds (Krumhansl, 1990; p.221). However, some sense of the initial key was maintained throughout the modulating passage. Since modulation is common in Western music, this ability to switch rapidly between schemas might pertain only to key-related schemas. One might imagine that switching, say, from a Western string quartet to Beijing opera would take longer -- although perhaps not very long in absolute duration. Bi-lingual speakers differ in their abilities to switch rapidly between different languages. But this skill appears to be related to how often speakers must change language in their daily life.

What cues signal the listener to switch schemas? Two plausible sources of cues for schema switching can be identified: auditory and non-auditory. One source might be obvious and persistent failures of expectation. Once again, switching between two languages is instructive. If a person has been conversing in French, then the failure of an utterance to conform to the schematic expectations for French ought to lead to a re-evaluation of the language context, and so precipitate switching to a different language schema. Similarly, the failure of pitch-, rhythm-, timbre- or other related expectations might be expected to instigate a search for a more appropriate schema.

A second source of pertinent cues can be found externally to the sounds themselves. For example, seeing five brass players on a concert stage will already evoke certain associations and expectations. If the players were dressed in dark evening suits, even more specific expectations might arise. Conversely, if the players were dressed in military uniforms, or if the players were dressed informally and standing on a New Orleans street, the expectations

would differ. There are innumerable visual and other environmental cues that presumably pre-dispose the listener to invoke a particular musical schema.

The auditory and non-auditory cues the provoke schema switching might also provide plausible cues through which new schemas are created. The persistent failure of expectations might well raise the alarm that a novel cognitive environment has been encountered and that the listener's existing pallet of schemas is inadequate. An interesting consequence of this view is that it should be difficult to form a new schema when the new context differs only slight from an already established schema. Once again, language provides a useful analogy. Native English speakers who learn a latinate language, often encounter difficulty learning a second latinate language. For example, a non-fluent knowledge of Spanish may interfere with the ability to learn Italian. Italian vocabulary and grammar may begin to interfere retroactively with one's Spanish abilities. The difficulty appears to be the failure, from an English speaker's perspective, to sufficiently distinguish Italian from Spanish. This confusion appears to be reflected in neurological studies. It is often the case that cortical areas associated with a native language are segregated from cortical areas associated with an acquired second language. However, a third acquired language will often share cortical regions associated with the second acquired language. In this case, the weak cognitive barrier between schemas is reflected in an apparently weak neurophysiological barrier.

Whatever form these barriers take, they are clearly important in order to maintain the modular structure of auditory schemas. As we noted earlier, these cognitive

barriers allow a listener to be surprised by events that in one schema are common, but in another schema are uncommon. While a modern listener might be quite familiar with jazz, this same listener might well find a moment of syncopation in a Renaissance motet to be somewhat "shocking." Such experiences imply that relatively strong barriers exist between schemas. Indeed, in Castellano, Bharucha and Krumhansl (1984) it was found that American listeners did not carry over Western pitch expectations to the experience of listening to North Indian music [check this]. More research is clearly needed to determine the extent to which one musical schema can influence another.

Cross-Over

What happens when the events of the world straddle two different schemas? The apparent modularity of auditory schemas suggests that the boundaries between schemas provide musically fruitful opportunities for playing with listeners expectations. Many musically interesting "cross-overs" have arisen over the years. For the author, one such distinctive experience can be found in *Bach Meets Cape Breton* recorded by David Greenberg and the group *Puirt a Baroque*. Greenberg received classical training as a baroque violin specialist, but Greenberg is also an accomplished Cape Breton-style fiddler. In recording traditional baroque dance suites, Greenberg shifts easily between conventional art-music interpretations and traditional fiddling. A "gigue" by Bach will morph into a "jig." One has a palpable sense of connections being made between two formerly discrete musical schemas. A listener

begins to imagine a continuum between courtly baroque dances and 18th century folk dances.

From a musical point of view, stylistic and genre distinctions contribute to the wealth and variety of musical experience. As we have seen, experienced listeners probably form different stylistic schemas for renaissance and rock music, between blue grass and bebop. As psychological constructs, however, genres exist as encapsulated expectation-related knowledge. The knowledge is modularized in separate schemas as the brain's way of preventing past experiences from being over-generalized to inappropriate contexts. When creating new styles or genres, musicians take advantage of the existing evolutionary cognitive machinery for protecting an organism from misapplying local information to other environments. The fact that the brain so readily brackets novel environments suggests that musicians have considerable latitude for creating new and unprecedented musics.

Schema Failures

Schemas can fail listeners in two ways. We may fail to apply the correct schema to a given listening situation, or our schema may be flawed in some way. We have already seen that listeners do not always learn the "right" principals of organization. Even though two genres of music may different in their underlying principals of organization, it is possible that listeners are incapable of distinguishing the two genres. Said another way, it is possible that attempts to create a new genre will fail, because the new

genre does not engender a significantly different set of expectations.

Alternatively, listeners may simply fail to gain sufficient exposure to bring about the creation of the new schema. Such failures are commonplace when listening to the music of an unfamiliar culture. However, such failures can also occur within one's culture. In Western music, an example can be found in the perception of atonal music. Krumhansl, Sandell, and Sergeant (1987) found that listeners to atonal pitch sequences divided into two groups. One group of listeners had internalized atonal conventions and judged as ill-fitting those pitches that had appeared recently. However, a second group of listeners continued to hear the sequences according to tonal expectations. The two groups were found to differ in musical background -- the former group being more highly trained. This implies that greater exposure would have benefitted the second group of listeners.

The experience of atonal listening is described in more detail below. However, Krumhansl and her colleagues found no evidence for a truly unique "atonal" way of listening. Rather, their results suggest that diatonic tonal hierarchies continued to be used by all listeners, but that some listeners systematically responded in a manner contrary to the tonal schema -- a sort of musical "reverse psychology" (see below).

REALIZED, THWARTED, MIXED, REVERSE, AND PARADOXICAL EXPECTATIONS

What happens when a listener's expectations prove correct? Conversely, what happens when a listener's expectations prove incorrect? In his book, *Emotion and Meaning in Music,* Leonard Meyer proposed the important hypothesis that expectations are intimately tied to emotional responses. In particular, Meyer suggested that thwarted expectations cause uneasiness or anxiety for listeners. For Meyer, "the frustration of expectation [is] the basis of the affective and the intellectual aesthetic response to music." (p.43).

Meyer argued for a sort of generalized emotion related to expectation. Contemporary empirical research supports this view in what has become known as the *primary affect* arising from expectation. However, the research further implies that different emotional responses are evoked depending on the nature of the expectation and its relationship to the actual outcome. With regard to primary affect, at least five conditions need to be distinguished: (1) when outcomes match the listener's expectation, (2) when outcomes conflict with the listener's expectation, (3) when some expectations are confirmed while others are simultaneously thwarted, (4) when a listener learns to expected the unexpected, and (5) when a listener experiences a single outcome as paradoxically both expected and unexpected.

1. Expectations Fulfilled

Expectations that are fulfilled represent stunning mental achievements. When a listener correctly anticipates that a dominant seventh chord will resolve to the tonic, this seemingly simple skill bears testament to millions of years of evolution that have shaped sensory and perceptual systems. Brains have evolved explicitly to make such accurate predictions possible.

Since the purpose of expectation is to anticipate events in the environment, accurate expectations may be deemed "successess" while inaccurate expectations constitute "failures." One might well imagine that expectational failures would engender stress, whereas expectational successes would engender some feeling of satisfaction or enjoyment. This simple principal carries significant repercussions for understanding the exposure effect -- discussed earlier in connection with tonality. Recall that listeners exhibit a preference for the most commonly occurring stimulus.

In the absence of any other evidence, it is reasonable for an experimental subject to predict that the next stimulus will be the most commonly experienced stimulus in the experiment. The pleasure or preference reported by subjects in these experiments may not be directly attributable to exposure. An alternative interpretation is that subjects experienced a moment of phenomenal pleasure because the most commonly encountered stimulus had unconsciously been predicted. In short, the exposure effect might itself be an artifact of positive affect evoked by accurate anticipation.

What, we may ask, is the consequence of getting things right? Part of a listener's response will depend on the

associated consequence of the anticipated state. For example, a listener might predict that the cracking of a branch overhead will be followed by the thud of something hitting the ground. In this case, our expectation might provoke a motor behavior in which we step out of the way, or look up. In other cases, forming accurate expectations might suppress a response. For example, in a darkened room we may hear the sound of something moving across the floor. Our penchant to become fearful may be suppressed by the accurate prediction of the reassuring sound of one's cat meowing.

In the case of music, the consequences of our predictions are less onerous than the sound experiences our ancestors might have had in an unforgiving pleistocene environment. Nevertheless, remember that anticipating events is one of the things brains are built for. We cannot "turn off" our tendency to anticipate. Since expectations have strong survival value, it is not farfetched to suppose that the brain itself provides reward mechanisms for accurate predictions. That is, it is possible that listeners experience a small positively valenced emotional charge when expectations are fulfilled. In other words, it may not be familiarity *per se* that evokes preference; instead preferences may arise from successful expectation.

On the other hand, repetitive sounds can lead to boredom. There is no challenge in predicting that the swishing sound of an electric fan will be followed by more swishing sounds. Habituation is nature's way of getting an organism to ignore stimuli that carry no information.

How do we reconcile the preference for familiar stimuli with the experience of boredom? Note that all of the

experiments that show people prefer familiar stimuli have been carried out using sparse stimuli. The amount of repetition used in these experiments was small, so no habituation would be expected.

When our surroundings become highly predictable, we become bored. The behavioral consequences of such situations is a lowering of arousal levels, a reduced attentiveness, and often a tendency to become drowsy and perhaps fall asleep. Since periodic sleep is biologically necessary, what better place to sleep than in an environment that is utterly banal and predictable. There are good reasons to be reassured by familiar surroundings. There are also good reasons why we might show little interest in such surroundings.

It bears reminding that habituation is not possible with all stimuli. For example, people do not habituate to painful stimuli. When an especially loud sound is continuously repeated, for example, the effect will be one of annoyance rather than boredom. In short, not all highly expected stimuli will evoke reassurance.

2. Expectations Thwarted

Incorrect expectations cause stress. In ordinary life, people who experience constant and unpredictable change are known to suffer from high levels of stress. It is likely the case that thwarted expectations engender a release of cortisol -- a stress hormone. From an evolutionary perspective, failing to predict the environment increases risk. It reduces an organism's ability to take advantage of opportunities, or prepare for possible dangers. Thwarted expectations might be expected to raise arousal levels,

heighten attention, and encourage reappraisal and learning. Indeed, viewing unexpected stimuli causes galvanic skin responses consistent with increased arousal.

Expectations do not go away simply because reality doesn't comform to them. Three sorts of responses might be imagined in response to thwarted expectations. In the first case, the expectation for a specific outcome may be retained, and the listener continues to expect a given outcome, even though it hasn't yet happened. If the expectation is finally fulfilled, then the principal aesthetic or emotional effect will relate to *delay.* The stress of uncertainty will be short-lived and the listener is likely to experience some measure or "relief" of the "I-knew-it-all-along" sort.

Another possibility is that the listener has applied the wrong expectation to the passage. That is, the listener may have misapprehended the context. For example, a listener might have the expectation that a tonic (I) chord is not typically followed by a bVII chord. However, if a third chord (IV) ensues, then the listener might reconceive of the passage: if the first chord is regarded as a dominant (V) chord, then the passage because a (more probably) V-IV-I progression. In other words, a thwarted expectation might engender a reappraisal of the context to ensure that the correct schema is being applied.

A final possibility is that the predictive failure is total. That is, the events cannot be attributable to a delayed fulfillment or a misapprehended context. The listener is unable to reconcile the actual events with any existing perceptual schema they may have. In this case, the listener will experience a relatively high degree of stress

and discomfort. Of course, the usual ongoing learning will continue, so unconscious processes will code the event and update or create a possible new schema to account for such experiences in the future.

Consider, by way of example, a Western listener who has had little or no experience with atonal music. For this listener, sequences of notes will systematically fail to conform with any existing schema. The music is likely to be experienced as stressful and uncomfortable. But with repeated exposure, the listener will slowly develop the kinds of expectations shown by experienced atonal listeners. With this new schema in place, subsequent listening experiences will be significantly less stressful.

3. Mixed Expectations

Expectations rely on some underlying mental representation. Listeners expect something concrete -- like a particular pitch, or harmony, or tone color. In the case of music the extant experimental literature implies that listeners typically maintain several concurrent musical representations. This suggests that that a given musical event might be surprising from the perspective of one representation, but entirely expected from the perspective of another representation. A possible musical example of mixed representations leading to mixed outcomes is evident in Figure 29. The passage is taken from a flute sonata by Benedetto Marcello. A sequence in the upper (flute) part is repeated three times. In the first and second sequences 4-3 suspensions correspond to the high point in the phrase. However, in the third instance of the sequence,

the suspension drops down an octave (arrow) from where it might have been expected.

Figure 29

Fig. 29. Excerpt from Marcello's *Sonata* in A minor for flute, measures 46-54. Three instances of a sequence are shown. In the third instance, the pitches C5 and B4 are an ocatve lower than would be expected. However, the harmonic sequence is preserved.

The octave displacement here would be surprising if the passage is mentally represented using pitch contours or intervals. However, the final three notes would not be surprising if the passage is mentally represented using pitch-classes, or "pitch-class contour". Moreover, these changed notes still preserve the underlying harmonic sequence. The *continuo* part harmonizes each sequence as a *V-of* harmony ending in a 4-3 suspension. In other words, the final three notes evoke "surprise" for pitch, contour, and interval representations, whereas the notes are entirely expected for pitch-class, pitch-class contour, and harmonic representations.

4. Reverse Psychology: Expecting the Unexpected

Another form of expectation arises when listeners learn to expect the unexpected. In a famous passage outlining his method of composing with twelve tones, Schoenberg claimed that repeating a pitch has a tendency to raise the tone to the status of the tonic. Given his avowed aesthetic

goal to avoid tonality, Schoenberg proposed a remarkably simple system of constructing a tone-row where all twelve pitch-classes are sounded one after another. In effect, Schoenberg advocated creating music where the aggregate distribution of pitch-classes shows a "flat" or uniform distribution. Notice that this compositional approach is very much consistent with the view that the perception of pitch stability tends to be related to an unequal pitch-class distribution where one or another pitch becomes more predictable.

Of course, tonal implications are hard to eliminate. As we have seen, playing just a single tone is apt to evoke a sense of tonic for most listeners. In the construction of a tone row, a composer might well choose ensuing pitches so that they tend to erase any latent tonal implications. For example, beginning with the pitch `C', an ensuing `G' would tend to reinforce a C-major key implication; an ensuing `C#' or `F#' would tend to contradict the tendency to assume a C-major key context.

Huron and von Hippel (2000) carried out a detailed study of the construction of 12-tone rows from the classic "Second" Viennese school composers: Arnold Schoenberg, Anton Webern, and Alban Berg. Using some 80 twelve-tone rows, Huron and von Hippel examined the moment-to-moment key implications using the Krumhansl and Schmuckler key-estimation algorithm. The moment-to-moment unfolding of the tone rows were shown to exhibit strong contra-tonal organizations. By way of illustration, consider the first four pitches in Schoenberg's tone-row for Opus 27, No. 3: G, F#, D, and E. Given these four notes, there are eight possible choices for the ensuing (fifth) pitch-class. Table 4 shows the maximum Krumhansl and

Schmuckler key correlations that arise for each of the eight possible continuations for the fifth pitch-class. For example, continuing the row with pitch-class `A' causes a high maximum key correlation (r=+0.81 for D major), whereas continuing the row with `F' produces a low maximum key correlation (r=+0.43 also for D major).

Table 4

Initial Row	Possible Continuation	Maximum Key Correlation
G, F#, D, E	**C**	+0.64
G, F#, D, E	**C#**	+0.50
G, F#, D, E	**D#**	+0.47
G, F#, D, E	**F**	+0.43
G, F#, D, E	**G#**	+0.46
G, F#, D, E	**A**	+0.81
G, F#, D, E	**A#**	+0.55
G, F#, D, E	**B**	+0.79

If Schoenberg wished to circumvent this key implication, the best (lowest) key correlation would arise for the pitch **F** -- according to the Krumhansl and Schmuckler algorithm. The actual fifth pitch selected by Schoenberg is indeed **F**. In Huron and von Hippel, this contra-tonal tendency is evident throughout the twelve-tone rows used by these Viennese composers.

In another study of twelve-tone rows, Krumhansl, Sandell and Sergeant (1987) asked listeners to judge the "goodness" of various probe tones at successive points in a twelve-tone row. Interestingly, Krumhansl *et al*'s listeners divided into two distinct groups. Some listeners tended to rate "highly" tones which tended to reinforce some latent possible key. That is, the most highly rated tones tended to be those which maximized the aggregate correlation for the passage with the Krumhansl and Kessler key profiles. The

second group of listeners responded in a completely opposite fashion. That is, they rated most highly those continuation pitches that minimized the aggregate correlation for the passage with the Krumhansl and Kessler key profiles. In other words, this second group of listeners thought the most appropriate pitch continuations are those that create the most contra-tonal effect.

Fascinatingly, Krumhansl and her colleagues found that the two groups differed in their musical experience. The group that rated highly the most atonal continuations were the more musically experienced or trained listeners. This suggests that these listeners had internalized the contra-tonal organization underlying this music and were able to form expectations that correspond both with the aesthetic goal, and with the pitch-related statistics exhibited by the music. In other words, the bifurcation in listening strategies reflected the combination of the bifurcation of composing strategies, and the experience of the listeners.

The phenomenon of "expecting the unexpected" has repercussions for understanding musical enjoyment. Earlier it was claimed that the exposure effect may simply be an artifact of a postive affect evoked by accurate anticipation of stimuli. If this is the case, then the frequency of occurrence of a stimulus does not, by itself, engender a positive affect. The more pertinent issue is the degree of predictability. To the extent that knowledgeable listeners are better able to predict the behavior of 12-tone music, then it should not be unexpected that knowledgeable listeners might enjoy 12-tone music more than other listeners.

On the other hand, it might be noted that the expectations of knowledgeable listeners when encountering 12-tone music are rather vague. Knowledgeable listeners have a higher than chance ability to predict which pitch-classes are *unlikely* to occur next. But there may very well be a difference between knowing which two or three stimuli are most likely to occur next, and which two or three stimuli are least likely to occur next. It may be that expectation-evoked pleasure arises foremost when an expected stimulus is realized, not when an unexpected stimulus in not realized. It is possible that this hypothetical asymmetry limits the expectation-related pleasure that can arise from listening to 12-tone music.

5. Paradoxical Expectations

The famed philosopher, Ludwig Wittgenstein, described a paradox that has bothered generations of music scholars. How is it possible, asked Wittgenstein, for a listener to be surprised by a work whose familiarity means that it can hold no surprises? (Wittgenstein, 1966). Jay Dowling and Dane Harwood proposed that the paradox might be resolved by distinguishing conscious from subconscious listening experiences (Dowling & Harwood, 1986; p.200). Dowling and Harwood proposed that we hear familiar pieces against the background of schematic norms for various styles and genres.

Jamshed Bharucha (1987) proposed a more precise distinction between two kinds of expectations: *schematic* and *veridical*. Schematic expectations arise from a lifetime of music listening. Schematic expectations arise without conscious thought and cannot be easily suppressed. "Even

when a given piece has been heard often enough to be familiar, it cannot completely override the generic, automatic expectations. Surprises in a new piece thus continue to have a surprising quality because they are heard as surprises relative to these irrepressible expectations." (Bharucha, 1994; pp.215-216) But Bharucha goes on to say that schematic expectations alone cannot account for common listening experiences: "If the surprises in a new piece continue to be surprises even after repeated hearing, the piece would never sound familiar." (p.216). Accordingly, two systems related to expectation must exist.

The Tenacity of Schematic Expectations

If a listener knows exactly what is about to happen, then surely, if the coming event contradicts the normal schematic expectations, then these schematic expectations can be ignored or suppressed. Not so. In an experiment by Bharucha and Stoeckig (1989), they pitted schematic and veridical expectations against each other with revealling results.

Once again, the task was for listeners to identify whether the target chord was in-tune or out-of-tune. But the stimuli were presented twice in succession before the listener responded. For example, in the "unexpected" condition, a listener might hear a C-major chord followed by an F#-major chord, followed by a pause, followed by a repetition of the C and F# chords. When the listener responded, the listener already knew what chord to expect. That is, the listener's veridical expectation was for the F#-major chord -- even though this progression violates the common

schematic expectation for a more closely related chord. In half of the trials, the last chord was mistuned. Despite the fore-knowledge of what chord to expect, the schematically expected chords were still processed more quickly than the schematically unexpected chords. That is, the schematic expectations remained influential, even when the listener knew exactly what was coming.

The tenacity of schematic expectations provides a plausible explanation for why, for example, a deceptive cadence will still sound somehow "deceptive" even though the listener fully expects it.

Meyer proposed that it is possible for listeners to apply the wrong schema: "the same physical stimulus may call forth different tendencies in different stylistic contexts ... For example, a modal cadential progression will arouse one set of expectations in the musical style of the sixteenth century and quite another in the style of the nineteenth century." (Meyer, 1956; p.30)

IV. PSYCHOLOGICAL CONSEQUENCES OF EXPECTATIONS

As we have noted, the ability to anticipate future events is important for survival. Minds are "wired" for expectation. However, from the subjective or phenomenological point of view the most important aspects of expectation are the feelings they are capable of evoking. What happens in the future matters, so it should not be surprising that how the

future unfolds has a direct effect on how we feel. In particular, music scholars have long noted that music-related expectations are capable of evoking emotional experiences.

In considering expectations, four different types of emotional responses can be distinguished. Two types of emotional responses occur prior to the event and so might be dubbed *pre-outcome* responses; two further types of responses are associated with the final outcome and might be dubbed *post-outcome* responses.

1. Imaginative Response

The first type of emotional response arises from imagining some future outcome. Imagining an outcome allows us to take some vicarious pleasure (or displeasure) -- as though the outcome has already happened. We may choose to work overtime because we can imagine the embarrassment of having to tell the boss that a project remains incomplete. We may be motivated to undertake a difficult journey by imagining the pleasure of being reunited with a loved one. This *imaginative response* is important in behavioral motivation. Through day-dreaming, it is possible to make future outcomes emotionally palpable. In turn, these feelings motivate changes in behavior that can increase the likelihood of a favorable outcome.

Neurological evidence for such an imaginative response is reported by Damasio (1994), who has described a neurological condition in which patients fail to anticipate the feelings associated with possible future outcomes. In one celebrated case, Damasio described a patient who was capable of feeling negative or positive emotions after an

outcome had occurred, but was unable to "preview" the feelings that would arise if a negative outcome was immanent. Although Damasio's patient was intellectually aware that a negative outcome was likely, he failed to take steps to avoid the negative outcome because, prior to the outcome, the future negative feelings were not palpable and did not seem to matter. Damasio's work establishes that it is not simply the case that people think about future outcomes; when imagining these outcomes, we are also capable of feeling a muted version of the pertinent emotion. We don't simply *think* about the future possibilities; we *feel* future possibilities.

The *imaginative response* provides the psychological foundation for deferred gratification. Feelings that arise through the imagination help individuals to foresake immediate pleasures in order to achieve a greater pleasure later.

2. Tension Response

The second type of pre-outcome emotional response arises due to uncertainty in high-stakes situations. Sometimes outcomes are utterly certain and have litte consequence. In other cases, we may have little idea about what is about to happen. If one or more of the possible outcomes involves a high stake (something very good or very bad), then we will tend to be more alert as the moment approaches when the outcome will be made known. Specifically, our physiological arousal level will be high. Heart rate and blood pressure will typically increase, breathing will become deeper and more rapid, perspiration will increase, and muscules will respond faster. These and other physiological changes help

us to react more quickly, and to attend and perceive more accurately. However, these changes are also associated with stress.

This type of pre-outcome response might be called *tension responses*. The stress or tension is proportional to the amount of uncertainty, and to the difference in magnitudes between the best and worst outcomes. The difference in magnitude is important. For example, a lottery winner may be relatively unconcerned as to whether the final prize is $68 million or $74 million. A large degree of uncertainty may surround the ultimate outcome, but the actual resolution of the outcome may be perceived as inconsequential. The tension response is independent of whether the anticipated outcome is positive or negative. Thus, when sentencing a convicted shop-lifter, the choice of prison term may be between 95 and 110 days. While there may exist a high degree of uncertainty about the precise sentence, the tension response may be muted because the difference between the outcomes is relatively small.

The tension response is also influenced by the elapsed time before the outcome is known. As the anticipated moment of outcome approaches, the tension increases. That is, tension is inversely proportional to the estimated remaining time to the onset of the outcome. There are good reasons why tension should increase as the outcome approaches. High arousal and attention are most needed at the point where one must respond to the outcome.

Simon Durrant has noted that, in general, organisms should try to avoid situations of high uncertainty. High uncertainty requires arousal and vigilance, both of which incur an energy cost. Consequently, it would be adapative

for an organism to experience high tension responses as unpleasant. That is, even if only positive outcomes are possible, high uncertainty will lead to an unpleasant stress.

By way of summary, it is proposed that the tension response is shaped by three factors: (1) the degree of uncertainty, (2) the estimated amount of time before the outcome is realized, and (3) the range separating the most positive and most negative outcome (that is, the "stakes" of the outcome).

3. Outcome Response

Two further types of emotional responses occur only once the outcome is known. The most obvious of these emotions relates to the pleasantness or unpleasantness of the outcome, such as the "fear" of encountering a snake, the "sadness" of receiving a poor grade, or the "joy" of giving birth. We might refer to these state-related emotions as the *outcome response*. These types of emotions have been the subject of extensive research and will be addressed at length in a later chapter.

Here we need only note that positive and negative emotions act as behavioral reinforcements. The pain caused by biting your tongue teaches you to chew carefully and avoid tissue damage. Bad tastes and bad smells reinforce the aversion to ingesting unhealthy foods. The pleasure caused by engaging in sex encourages procreation. The enjoyment of playing with our children, encourages parental investment and nurturing. Positive emotions encourage us to seek out states that increase our adaptive fitness. Negative emotions encourage us to avoid maladaptive states.

4. Prediction Response

Recall that an expected stimulus is more accurately perceived when it is predictable. Accurate predictions help an organism to prepare to sidestep dangers and take advantage of opportunities. Since accurate predictions are of real benefit to an organism, it would be reasonable for psychological rewards and punishments to arise in response solely to the accuracy of the expectation. Following a snow storm, for example, I might predict that I will slip and fall on the sidewalk. In the event that I actually fall, the outcome will feel unpleasant, but the experience will be mixed with a certain satisfaction at having correctly anticipated the outcome. This fourth type of expection-related emotion might be dubbed the *prediction response*.

Psychological evidence in support of a *prediction response* is found in the work of Mandler (1975). The response is considered so important in the extant literature on expectation, that it is commonly referred to as the *primary affect* related to emotion (Olson, Roese & Zanna, 1996). [5] Confirmation of expected outcomes generally induces a positive emotional response even if the expected outcome is bad. It is as though brains know not to shoot the messenger: accurate expectations are to be valued (and rewarded) even when the news is not good. That is, a person might experience a positive *prediction response* and a negative *outcome response* at the same time.

In summary, we have distinguished four different types of expectation-related emotions. Each type serves a different biological function. The purpose of imaginative responses is

to motivate an organism to behave in ways that may maximize future benefits. The purpose of the tension response is to tailor arousal and attention to match the level of uncertainty and importance of the outcome. The purpose of the outcome response is the often-noted goal of all emotions: to provide positive and negative reinforcements related to the biological value of different states. The purpose of the prediction response is to provide positive or negative reinforcements related to forming accurate expectations. All of these goals are biologically valuable.

Response type	Epoch	Biological Function
imaginative response	pre-outcome	future-oriented behavioral motivation
tension response	pre-outcome	optimum arousal & attention in preparation for possible events
outcome response	post-outcome	negative/positive reinforcements related to specific states
prediction response	post-outcome	negative/positive reinforcement to form accurate expectations

Informally, we might characterize the "feeling" components to these responses by posing four questions:

1. What do you think might happen, and how do you feel about that?
2. Are you ready for what's about to happen?
3. How do you feel about how things have turned out?
4. Did you place a good bet?

Expecting *What* and *When*

As noted earlier, predicting a future event actually entails two predictions: the *what* and the *when*. The predictability of the *what* and *when* can be entirely independent. In

musical rhythms, for example, listeners can form a strong expectation that some sound will happen at a particular moment, even though they have little inkling of *what* sound will occur. In other circumstances, the listener will have a good idea of *what* to expect, but will be left wondering *when* the sound will happen.

As in the case of accurately predicting *what* will happen, accurately predicting *when* an event occurs will facilitate perception. In the work of Jones *et al* discussed earlier, we saw how listeners are able to more accurately process a sound when it occurs at a predictable rhythmic moment.

Listeners often claim that an unpleasant sound will seem "abrupt" sounds. Webster's dictionary provides two pertinent definitions for *abrupt*: "1. occurring without warning, UNEXPECTED" and "2. rising or dropping sharply as if broken off". Both of these definitions are pertinent to the experience of sound. An abrupt sound is often simply a sound that is unexpected. In addition, an abrupt sound may have an especially rapid onset. The sound of a cat starting to purr has a much slower onset than the sound of a bursting balloon. The slower acoustical onset provides the listener with slightly more time to prepare for the sound before it reaches maximum amplitude. That is, a slower sound onset provides a split second in which the auditory system can prepare (predict) for what is likely to happen next. A sound can be "abrupt" both because it occurs at an unpredictable time, and because the sound itself has a low predictability.

The Poetry of Expectation

The *what* and *when* components of expectation can be clearly seen in the case of poetry. Two features of poetry are known to appeal to listeners: a *rhyme scheme* and a regular *meter* or rhythm. Consider, by way of example, the following stanza:

> Life's not so short
> I care to keep
> The unhappy days;
> I choose to *sleep*. [6]

The poem exhibits a duple meter with two iambic beats in each line. Consider the listener's expectation at the moment prior to the last word (*sleep*). By establishing this regular meter, listeners expect the final syllable to coincide with the second beat. That is, the meter establishes a high expectation of *when* the final syllable will occur.

In addition, listeners will expect the final vowel to rhyme with the "ay" of "days" (or the "ee" of "keep"). That is, the poem provides listeners with helpful clues of the *what* of the final syllable. The rhyme scheme directly facilitates the perception of the final vowel.

There are good reasons why people might prefer poems that have a rhyme scheme and regular meter. These structures make the sounds more predictable, and so easier to perceive and process. But more importantly, the fact that listeners are able to accurately anticipate future events means that the auditory system evokes a positively valenced *prediction response*. Unconsciously, the brain is rewarding itself for doing such a good job of anticipating stimuli.

Predictability and Boredom

As noted earlier, high predictability can also lead to boredom. In highly predictable environments, the *tension response* falls to zero. No preparation is needed in anticipation of ensuing stimuli. There is no need to be attentive or aroused, and consequently minimal stress is evoked. The behavioral consequences are boredom and sleep.

When an environment is highly predictable (utterly lacking in novelty), the tendency is for an organism to become sleepy. Highly predictable environments are typically safe, and so nature takes advantage of the opportunity to reduce arousal levels and conserve energy.

Musical Applications

The preceding model of expectation can be applied to music in a number of ways. A useful exercise is to consider common conventions found in Western art music. For example, embellishments such as anticipations and suspensions have often been regarded by music theorists to involve expectation-related nuances. Below, we analyze four common types of embellishments: the anticipation, the suspension, the passing tone, and the appoggiatura.

In analyzing these embellishments, we will consider the predictive-, tension-, and outcome-related responses arising at each moment as the embellishment is approached and resolved. Due to the complexity involved, we will not consider imaginative responses. [7] In addition,

we will need to analyze separate the *what* and the *when* dimensions of expectation.

The Anticipation

By way of example, consider the *anticipation* illustrated in Figure 30. Here the anticipation occurs as part of an authentic *V-I* cadence with the final tonic pitch anticipated. The numbers identify three moments that we will analyze separately. The moments can be designated the (1) pre-anticipation, (2) anticipation, and (3) post-anticipation moments.

(1) Consider first the pre-anticipation moment.

Figure 30a

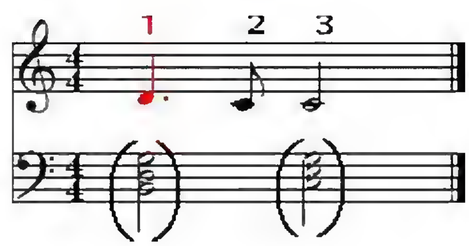


Fig. 30a. An example of an anticipation in a cadential *V-I* context.

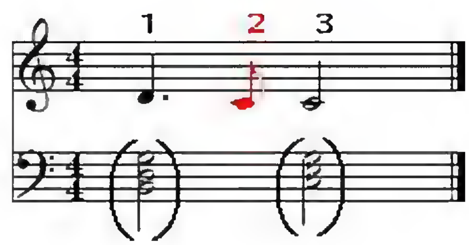
Outcome response: With an already established key context, the listener hears a dominant chord. The chord itself is the "outcome" of preceding expectations. As an outcome, we need to consider its response valence. Since the chord is a simple major sonority, it exhibits a low degree of sensory dissonance and so will tend to evoke a relatively positive valence.

Tension response: At the same time, musicians would note that the dominant function would normally be considered "dissonant" insofar as it needs resolution. This way of speaking can be re-interpreted in terms of the *tension response*. We would note that the *V* chord has a low probability of being followed by silence (i.e., it is unsuitable for closure). Experienced listeners will have a strong expectation that some further sounds will occur. Moreover, the *V* chord has a high probability of being followed by a *I* chord and the supertonic has a similarly high probability of leading to the tonic. In short, the listener has a relatively good idea of what to expect next; there is little of the stress that comes with uncertainty. Consequently, the tension response has only a very small negative valence.

There is one aspect to the tension response, however, in which there is relatively higher uncertainty. This has to do with *when* a tonic chord might appear. Since the dominant chord occurs on the downbeat, one possible moment of occurrence would be the downbeat of the next measure. Another possibility, might be the third beat of the current measure.

(2) Consider now the moment when the anticipation note appears (C eighth-note).

Figure 30b



Outcome response: The first thing to note is that the sonority is now more dissonant. That is, the *outcome response* has a comparatively negative valence.

Prediction response: Since the previous moment lead the listener to make a prediction, we can now consider the successfulness of this prediction. The pitch of the anticipation was indeed the optimum prediction arising from the previous moment, so there is a predictive "reward" associated with the "what". That is, the *prediction response* is positively valenced. However, the timing of the onset for this note is very low. Recall that the third beat or the downbeat of the next measure were more likely moments for "when" for this event might occur.

By way of summary, at the moment when the anticipation appears, the outcome response is rather negative, while the prediction response is a mix of positive ("what") and negative ("when").

Next, consider the tension response associated with this moment.

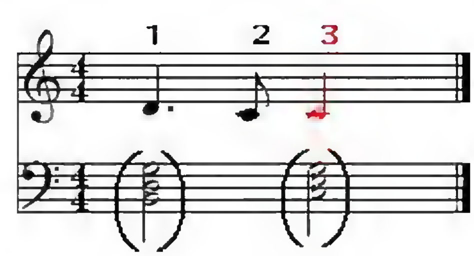
Tension response: Compared with the pre-anticipation sonority, the anticipation occurs on a surprisingly weak beat (the second half of the second beat). This very significantly raises the likelihood of an ensuing stimulus event occurring on the third beat. That is, the presence of the eighth-note significantly reduces the uncertainty as to whether the tonic chord will appear at beat three, or wait until the next measure. From the "when" point-of-view, the appearance of the anticipation greatly reduces uncertainty and so evokes a positively valenced tension response.

In addition, the pitch of the anticipation reduces the uncertainty concerning the ensuing "what". We know that listeners expect ensuing pitches to be close to current pitches, and that the closest possible pitch movement is unison repetition. Since the listener is already predicting that the *V* chord will be followed by a *I* chord, the appearance of the tonic pitch gives greater credibility to this prediction. That is, the presence of the anticipation lowers the uncertainty of "what" and so again contributes to a positively valenced tension response.

In the case of the anticipation, both the "what" and "when" components of listener uncertainty are reduced dramatically. Although the sonority is dissonant, and although the listener tended to predict a later occurrence of this pitch, the presence of the anticipation itself produces many psychologically positively valenced repercussions.

(3) Finally, the post-anticipation moment occurs.

Figure 30c



Outcome response: The outcome response is highly positive: the chord has low sensory dissonance.

Prediction response: The listener's confident prediction of this moment is realized, and so there is a high positively valenced prediction response.

Tension response: The closure associated with this moment creates a highly certain expectation that the current moment will be sustained for two or more beats, and perhaps followed by silence. That is, the tension response is also positively valenced since both the "what" and "when" following this moment are highly predictable.

Before leaving the anticipation, consider the variant passage shown in the figure below. Here, the duration of the anticipation has been increased to a quarter-note. Two important differences distinguish this case from the previous one. First, by falling on a more predictable beat, it reduces the likelihood of something happening on beat three. That is, the dotted-quarter/eighth of the original example makes it more certain that something will happen on beat three. In effect, decreasing the duration of the anticipation renders it more effective in helping the listener predict the "when" of the ensuing event.

Figure 31

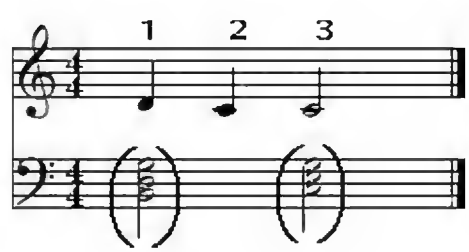


Fig. 31. A variant anticipation in which the duration of the anticipated note is extended. (See discussion in text.)

The second difference is that having the anticipation occur on beat two rather than the second half of beat three makes the anticipation note itself more predictable. In

effect, there is a trade-off between the predictability of the anticipation moment and the post-anticipation moment.

The observations made above concerning the anticipation are summarized in the following table. Responses colored in red indicate a negatively valenced response, whereas responses colored in blue indicate a positively valenced response.

Summary Expectation Analysis of Anticipation

	Outcome	Predictive	Tension
pre-anticipation	consonant	-	low tension; strong expection of the ensuing resolving pitch
anticipation	dissonant	high predictive success for pitch; low predictive success for timing	extremely low tension; nearly certain of ensuing resolving pitch; in a sense, the current pitch is the resolution of the previous expectation and so early outcome further reduces the tension
post-anticipation	consonant	extremely high predictive success	-

More than other embellishments, anticipations are more likely to occur near a cadence, and therefore arise in situations that are more predictable.

The Suspension

Figure 32 shows a typical 4-3 suspension. The suspension occurs as part of tonic-dominant progression in which the movement of the tonic pitch (F) to the leading-tone (E) is delayed. The numbers identify the (1) pre-suspension, (2) suspension, and (3) post-suspension moments.

(1) Consider first the pre-suspension moment.

Figure 32a

Outcome response: With an already established key context, the listener hears a tonic chord (in F major). The chord itself is the outcome of preceding expectations that we needn't consider. The chord is a simple major sonority with low sensory dissonance, and therefore will tend to evoke a positive valence.

Tension response: As a *I* chord, it is quite stable and so may evoke no strong sense of continuation. Nevertheless, a number of possible continuations might be expected, including a good likelihood of being followed by a *V* chord. In addition, pitch proximity will tend to engender expectations that the pitch F is likely to be followed by a nearby pitch (F, G, E). An experienced listener will therefore have a reasonable intuition of what might occur next: there is relatively little of the stress that comes with uncertainty. Consequently, the tension response exhibits a relatively small negative valence. As in the case of the anticipation example, one source of tension is *when* the ensuing chord/event might appear.

(2) Consider now the moment when the suspended sonority appears.

Figure 32b

Outcome response: The sonority is now more dissonant, so the outcome response has a comparatively negative valence.

Prediction response: The suspended note (F) has a high likelihood of following from the previous sonority. Similarly, the dominant chord is likely to follow from the previous tonic. This implies that a listener should typically experience a predictive reward associated with the "what". The combination of the expected pitch and the expected chord is probably less predicted. Nevertheless, the outcome is reasonably common and not unusual, so one would expect that the prediction response would be positively valenced for most experienced listeners. With regard to the "when", the suspended sonority falls on a highly predictable beat. It might have occurred a quarter-duration earlier, or perhaps a half-duration later, but the occurrence on beat three has a relatively high predictability. (The timing of the suspension here is more predictable than the timing of the anticipation seen in our earlier example.)

By way of summary, at the moment when the suspension appears, the outcome response is rather negative, while the prediction response is positive for both the "what" and the "when".

Tension response: The suspended pitch creates a very high expectation to move to the E. In other words, the "what" of the next moment is almost perfectly predicted. The "when" of the post-suspension moment is a little more uncertain. The resolution might occur on the next beat, or be delayed until the next major downbeat at the beginning of the next measure. However, relatively little uncertainty accompanies the "when". Only a couple of choices are likely. As in the case of the anticipation example, rather little uncertainty surrounds what will happen following the dissonant

moment. Consequently, the suspension evokes a positively valenced tension response.

(3) Finally, the post-anticipation moment occurs.

Figure 32c

Outcome response: the chord has low sensory dissonance and relatively high stability so the outcome response is highly positive.

Prediction response: The listener's confident prediction of this moment is realized, and so there is a high positively valenced prediction response.

Tension response: The closure associated with this moment creates a highly certain expectation that the current moment will be sustained for two or more beats, and perhaps followed by silence. That is, the tension response is also positively valenced since both the "what" and "when" following this moment are highly predictable.

Summary Expectation Analysis of Suspension

	Outcome	**Predictive**	**Tension**
pre-suspension	consonant	-	moderate to low tension; relatively strong expectation of the ensuing resolving pitch
suspension	dissonant	moderate predictive success due to proximity	very low tension; strong expectation of ensuing resolving pitch (via anchoring)
post-suspension	consonant	extremely high predictive success	very low tension; strong expectation of ensuing resolving pitch (via anchoring)
resolving	consonant	high predictive success	-

The Odd-ball Note

Given the preceding analyses a skeptical reader might conjecture that the introduction of any note would have a similar effect of reducing uncertainty -- and so produce positively valenced prediction and tension responses. As a control case, consider the concocted passage shown in Figure 33. This example shows a dominant-tonic progression with an "odd-ball" note interposed. A brief analysis follows.

(1) Consider first the pre-odd-ball moment.

Outcome response: With an already established key context, the listener hears a dominant chord with low sensory dissonance which tends to evoke a positively valenced outcome response.

Tension response: The dominant chord has a high probability of being followed by a tonic chord, and the supertonic pitch is likely to be followed by the tonic. Hence, the "what" component of the tension response has only a very weak negative valence. The "when" is slightly less certain. Plausible event onsets might occur on beat two, three, or the downbeat of the next measure.

Figure 32a

(2) Consider now the moment when the odd-ball note appears.

Outcome response: As with the anticipation and suspension, the sonority is now more dissonant, so the outcome response has a comparatively negative valence.

Prediction response: Both the pitch (A-flat) and the onset timing are poorly predicted, so the prediction response is highly negatively valenced. The A-flat does not belong to the key and so has a low probability of occurrence. In addition, the A-flat is remote in pitch from the preceding note, and is approached by the unlikely interval of a diminished fifth. The A-flat might be considered part of a dominant ninth chord -- a chord borrowed from the minor key. However, in general, the listener will receive little "reward" for predicting this event.

Tension response: The lowered sixth scale degree is typicaly anchored to the dominant pitch, so a reasonable prediction would be for the A-flat to be followed by G. Like the anticipation, the timing of the A-flat strongly implies that the next event should occur on beat three. Most experienced listeners would therefore confidently predict the occurrence of G4 on beat three. Both the "what" and "when" are highly predictable. Although the odd-ball note evokes negatively valenced outcome and prediction responses, it evokes a comparatively positive tension response.

Figure 32b

(3) Finally, consider the post-odd-ball moment.

Outcome response: the chord has low sensory dissonance and relatively high stability so the outcome response is highly positive.

Prediction response: The listener's confident prediction is clearly wrong. Both the "when" and the "what" fail to

conform to expectations. Only the fact that the chord is a tonic function was predicted. As a result, there is a highly negatively valenced prediction response.

Tension response: The tonic chord tends to evoke a sense of closure. However, the timing of the chord tends to reduce the closure effect.

Figure 32c

With only slight modifications our "odd-ball" example might be transformed into an appoggiatura (see Figure 33). An appoggiatura would have the A-flat resolving downward to the G on beat three. A more likely appoggiatura might employ an A-natural instead of the A-flat. But consider how this appoggiatura would evoke different expectation-related responses compared with the odd-ball passage. Both the odd-ball passage and the appoggiatura produce a dissonant moment, accompanied by a high expectation of the ensuing event. In the case of the appoggiatura, the subsequent resolution would conform to the expectation -- creating a positive prediction response in addition to the positive outcome response. However, in the odd-ball passage, the subsequent "resolution" fails to conform to expectations, hence evoking a negatively valenced prediction response.

Figure 33

Summary Expectation Analysis of Appoggiatura

	Outcome	**Predictive**	**Tension**
pre-	consonant -		moderate to low tension; relatively strong

appoggiatura			expectation of the ensuing resolving pitch
appoggiatura	dissonant	poor predictive success; surprising	low tension; strong expectation of ensuing resolving pitch (via anchoring)
post-appoggiatura	consonant	extremely high predictive success	-

Observations

What all four examples share in common is that the presence of the embellishment significantly increases the predictability of the ensuing sonority. In the case of the anticipation, appoggiatura and odd-ball, both the "what" and "when" of the subsequent sonority are made more certain. In the case of the suspension, the "when" is slightly less certain than the "what", but both remain high. The appoggiatura and the odd-ball produce a negatively valenced prediction response when the embellishment appears. However, the odd-ball passage also produces a negatively valenced prediction response at the "resolution" as well.

Looking at just the conventional embellishments -- the anticipation, suspension, and appoggiatura -- the presence of the embellishment creates a circumstance where uncertainty about the future is reduced. This is purchased at the cost of momentary dissonance. In other words, the negative valence evoked by sensory dissonance is balanced against the more positive valence of predictability. More precisely, the outcome valence at the time of the embellishment is made more negative, while the concurrent tension valence and the ensuing prediction valence (associated with the resolution) are both made more positive.

Misattribution and the Exposure Effect

Positive and negative emotions are important motivators that help organisms learn. Suppose I am mugged in a dark alley. I experience highly negative emotions whose purpose it to encourage me to avoid such situations in the future. But what, precisely, is the lesson I should learn? Should I learn to avoid dark alleys? Should I avoid encounters with other people? Should I avoid walking on concrete sidewalks? Should I avoid eating a sandwich for lunch? Once again, we are faced with the problem of induction: what general principal can one infer from finite observations? Moreover, since such highly emotionally-charged events tend to be rare, what can one reasonably learn from just one or two observations?

Nature addresses this problem by casting a very wide net. When we experience strong emotions, we tend to remember many details about the experience. A person trapped in a crashed automobile will tend to retain vivid memories of the crash site, the face of the ambulance attendant, and the music playing on the car radio. Research on *misattribution* has established that we tend to associate strong emotional experiences with all salient perceptual cues (time-of-day, facial features, manner of speaking, location, colors, etc.). Since the experience is highly charged, it is better to draw excessively broad conclusions (which have a better chance of catching a true cue) than to draw narrow lessons (that have a high chance of failing to capture a pertinent cue). In other words,

misattribution is a predictable consequence of the problem of induction.

Recall now our earlier discussion of the *exposure effect* -- the tendency for people to prefer stimuli that are expected. What could explain the origin of the exposure effect? The combination of the *prediction response* and *misattribution* allows us to offer a plausible explanation as to why commonly occurring stimuli would evoke a positively valenced emotional response.

Many outcomes are neither positively or negatively valenced. Yet if we predict such an outcome, a positively valenced prediction response ensues. In such circumstances, there is always the possibility that the positive prediction response will be misattributed to the stimulus that evoked the response. If state `A' is highly likely, and if we correctly predict the occurrence of state `A' on many occasions, then state `A' will tend to become associated with the positively valenced prediction response. With constant repetition, this misattribution tendency will be reinforced, and so we begin to misattribute the prediction response to the stimulus. To the extent that any frequently occurring stimulus will become more predictable, such frequently occurring stimuli will tend to accrue a positive emotional response. In effect, we now experience a positive *outcome response* for a previously neutral stimulus.

This phenomenon provides a plausible explanation for why the tonic pitch sounds "nicer" than other pitches. Similarly, this phenomenon provides a plausible explanation for why the "downbeat" is experienced as pleasurable. Viewed from the perspective of the outcome response, there is nothing

to favor one pitch over another. There is nothing inherently more pleasurable about D4 than E4. However, when these tones appear in a context that leads to certain expectations, the expected pitch will be experienced as evoking a more positive valence.

Of course listeners don't simply prefer the tonic pitch to all other pitches. The tonic pitch as a passing tone in a dominant harmony doesn't evoke nearly the pleasure of that same tonic pitch terminating a final cadence. But recall that cadences are more predictable, and that the occurrence of the tonic at a final cadence is very predictable. What we mean by "tonality" is a system of relationships that increase the predictability of certain sounds in certain contexts, that evoke both a highly positive prediction response, as well as a positively valenced outcome response that arises from misattribution of predictability with certain outcomes.

Predictable Music

All of the foregoing discussion leads to an obvious problem. If positively valenced responses arise from predictability, then wouldn't the most enjoyable music be utterly banal? Wouldn't the best sounding music be entirely predictable?

Some music does seem to conform to this implication. For example, "trance" music, "minimalism" and "drone" music do seem to exhibit highly predictable structures. However, there is plenty of music that isn't so obviously predictable, yet it is enjoyable.

One consideration is the phenomenon of habituation. Simply repeating the tonic pitch *ad infinitum* will lead to a

desensitization of the auditory response. Unless the stimulus is painful, organisms habituate to repeated stimuli.

However, the avoidance of habituation alone cannot explain the relative variety found in musical passages.

Response Interactions

Barbara Mellers and her colleagues have described an interesting phenomenon that might be regarded as an interaction between the *prediction response* and the *outcome response*. Consider the following experiment. Basketball players were asked to take shots from different positions around the court. Before each shot, the player was asked to estimate the likelihood of scoring a basket. Following each shot, the player was asked how good they feel. As you might expect, players are happiest when they make a shot and are unhappy when they miss a shot (i.e., positive and negative outcome responses). However, the degree of satisfaction/dissatisfaction is directly related to the player's expectation. The greatest unhappiness occurs when the player misses a shot that they judge to be "easy" and are happiest when they score a basket that is judged to have a low probability of success. In general, unexpected fortune or misfortune cause the greatest emotional responses. That is, low expectation amplifies the emotional response to the outcome.

This interaction has repercussions for how listeners experience sound. If a nominally unpleasant sound is not expected by the listener, then the sound will be perceived as even more unpleasant or annoying. Conversely, if a nominally pleasant sound is not expected by the listener, it

will tend to be perceived as more pleasant. A lengthy atonal passage is likely to lead the listener to expect further atonal sonorities. Terminating an atonal passage with a major chord will tend to heighten the pleasing effect. However, from the perspective of expectation, the most negative auditory experiences will occur when uncertainty is high, when what you expect doesn't occur, and when the outcome is unpleasant.

Emotional Effect of Delay

A potent component to the tension response is delay. To this point, we have talked about tension principally in relation to the *what* of expectation. However, an important component of the tension response arises from the *when* of expectation.

We noted earlier that the tension response increases as the estimated outcome moment approaches. If the outcome occurs prior to the anticipated time, then the tension response will fail to have reached its peak. On the other hand, if the outcome is late, the tension response will reach a peak and may be sustained as we wait for the outcome to materialize. In short, delay tends to magnify the tension response.

Another way of thinking about delay is that it increases uncertainty. As we have seen, an unexpected good outcome generally evokes a more positive response than if the outcome is fully expected. Similarly, an unexpected bad outcome is generally more disappointing than if the bad outcome is expected. However, these basic relationships are influenced by the effect of delay. Suppose that there is a strong likelihood of a good outcome. If a delay ensues,

the anticipation causes some doubt that the outcome will happen as expected. That is, delay provides opportunities to entertain doubts, and so delay has the effect of reducing the subjective probability. Consequently, a highly expected good outcome will evoke a greater positive response if preceded by a delay, since the delay, in effect, lowers the sense of certainty. Similarly, a highly expected bad outcome will evoke a less negative response if it ensues without delay. If a delay ensues, then the sense of inevitability will be tempered by thoughts that something might intervene to thwart the negative outcome.

As can be seen, the effect of delay is most marked when expectations are most certain. (We have the most to lose when we are virtually certain of a good outcome, and the most to gain when we are virtually certain of a bad outcome.) This means that the effect of delay in music will be greatest when applied to the most stereotypic, cliché, or predictable of events or passages.

Consider some of the most predictable aspects in Western music. The most predictable pitch is the tonic; the most predictable metric moment is the downbeat; the most predictable chord is the tonic chord; the most predictable diatonic pitch successions follow after the sixth and seventh scale degrees; phrase endings are among the most stereotypic (low information) musical moments in Western music.

The simplest and most direct form of delay is the *rallantando* or *ritard*. In most music, the greatest slowing occurs in the closing cadence of a work. Typically, this final cadence involves approaching the most predictable pitch, the most predictable chord, and the most predictable

metric moment. Cadences are especially ripe points for delaying tactics.

Nor is it the case that cadences are delayed only by slowing the tempo. The history of Western music is repleat with cadential delaying tactics. Indeed, many of the most seminal harmonic techniques originated as cadential interlopers. This includes the addition of the subdominant pitch in the creation of the dominant seventh chord, the suspension, the cadential 6-4, augmented sixth chords, the Neapolitan sixth, the pedal tone, the augmented triad, the dominant ninth and thirteenth chords, the pre-terminal false modulation, the interminable terminating *I* chord, and the deceptive cadence. The number of ways of delaying the musical end is legion. This same phenomenon is evident in film, where the denouement is often rendered in slow motion.

In the late Romantic period, composers such as Richard Wagner established the elided phrase in which cadence moments were avoided: the anticipated cadence would instead begin the ensuing phrase. In some ways, this delaying tactic reached its apex in the twentieth century with the advent of the fade-out. In Gustav Holst's *The Planets*, the fade-out is achieved mechanically, but with electronic sound recording fade-outs became routine. With the fade-out, music manages to delay closure indefinitely.

Figure 29

Fig. 29. An early example of a "fade-out" ending. "Neptune, the Mystic" from Gustav Holst's *The Planets* (1914). The passage is for female chorus.

> The performance instruction reads: "The chorus is to be placed in an adjoining room, the door of which is to be left open until the last bar of the piece, when it is to be slowly and silently closed." "This bar to be repeated until the sound is lost in the distance."

The effect of delay, and the interactions between the prediction response and the outcome response are summarized in Table 5. The outcome-related affect is appraised as having either a positive, negative, or neutral *valence*. Positve outcomes are associated with opportunity and pleasure; negative outcomes are associated with threat and displeasure. The primary affect (or "expectancy-accuracy affect") is either expected or unexpected.

Table 5

	Negative	**Neutral**	**Positive**
Expected	**Annoyance Resignation Sadness Crankiness**	**Boredom Stability Repose**	**Contentment Serenity Reassurance**
Unexpected	**Disappointment Startle Defense Disgust Anger**	**Interest Surprise**	**Delight Joy Surprise Wonder Astonishment**
Delayed	**Worry Foreboding Anxiety Tension Fear**	**Orienting Attention**	**Hope Craving Anticipation Savouring Relishing**

Table 5 provides different descriptive labels for negative-delayed and positive-delayed states. However, the

differences between these two states are probably less than suggested by these terms. The characteristic feeling evoked by both is a strong sense of uncertainty. We use the words "worry" and "hope" only to emphasize that the valence of the secondary affect.

One might think that the increased predictability of the embellishment tones runs contrary to Meller's work with the basket-ball players. Recall that the outcome response is especially high when the player makes a basket that is considered unlikely. This seems to suggest that uncertainty amplifies ... The important distinction is between the tension response and the prediction response. When a basket-ball player sinks a basket that was considered improbable, the tension response is muted: the player is reasonably certain that he will *not* score the basket. The prediction response is bad, but offset

In a study by John Sloboda from Keele University in England, music-lovers were asked to identify those musical passages they found most emotional. Sloboda (1991) found that "shivers down the spine" occurred most often in passages containing unexpected harmonies. Tears were most likely to be evoked by appoggiaturas or sequences of appogiaturas. Both of these experiences are consistent with the wonder, joy, and awe associated with unexpected positive outcomes. Examples might include an unexpected transposition, chromatic mediant chord, or sustained chord.

Note that the table includes the word "surprise" in both the unexpected/neutral cell and the unexpected/positive cell. In English, we don't have separate words to describe the distinction intended here. In the case of

unexpected/positive surprise, we associate the experience with wonder, awe, fascination, or amazement. By contrast, the unexpected/neutral surprise might be associated with experiences such as stupefy, stun, confound, bewilder or flabbergast. Even these terms are a bit too negative to be associated with a neutrally valenced appraisal.

In general, unexpected and delayed events raise arousal levels, whereas expected events frequently lower arousal levels. Habituation is the epitome of an expected event. When a habituated stimulus has a neutral valence -- that is, when it is appraised as having no consequence -- then there is a tendency toward boredom. Environments that have little consequence are safe environments, so it is not surprising that boredom also tends to be associated with sleepiness: sleep is an appropriate behavior in safe environments.

When expected events have a positive valence, there is also a tendency toward a low arousal state. However, the positive valence is apt to engage a person ("entertain" in the sense of maintaining attention), and so contentment and serenity are less likely to produce sleepiness as quickly as is the case for boredom. Nevertheless, the safety of the environment is ultimately likely to progress to sleep.

Emotional responses happen in response to global expectations as well as local (within the music) expectations. Linda Dusman (1994), for example, has noted that when members of a concert audience are introduced to a new work that defies straightforward comprehension, listeners are disappointed.

Notice that the above taxonomy accounts for all the seven basic emotions identified by Lewis (1995): sadness, anger, disgust, fear, interest, surprise, and joy.

Expectation Shapes Mental Representations

As we noted earlier, expectations imply some sort of mental representation. The *what* of expectation must be expressed in some language. Listeners will expect a pitch, or a pitch-class, or a scale degree, or an interval, a chord function, a combination of duration and scale degree, etc. We also saw evidence of a variety of different representations, and that listeners may use a combination of representations.

Ideally, the best mental representation would be the one (or ones) that most accurately reflect the organization of the real world. If the real world is organized according to scale degrees, then scale degree would be an appropriate mental representation. If the real world is organized according to a combination of (say) pitch contour, metric position, and diatonic interval, then the most appropriate mental representation would echo this organization.

But how is a brain to know which representation is the best? How can an auditory system learn to discard one representation in favor of another? Here expectation may play a defining and perhaps essential role. Expectation is an omnipresent mental process; brains are constantly anticipating the future. Moreover, we have seen that there is good evidence for a system of rewards and punishments that evaluates the accuracy or our unconscious predictions

about the world. A defective mental representation will necessarily lead to failures of prediction. Conversely, a mental representation that facilitates accurate predictions is likely to be retained. In effect, our mental representations are being perpetually tested by their ability to accurately predict ensuing events.

This claim carries an important implication. It suggests that the auditory system spontaneously is capable of generating several representations, from which the less successful can be eliminated. This in turn suggests that *competing concurrent representations* is the norm in mental functioning. It may well be that the brain begins by assuming a simple representation (such as absolute pitch). If the world is not organized in a manner consistent with absolute pitch (as in the persistent singing of `Happy Birthday' in different keys), then some other representation (such as interval or scale degree) will become more appropriate. However, any latent absolute pitch representation will be retained to the extent that it retains some value in predicting the future.

Expectation serves at least three functions: *motivation*, *preparation*, and *representation*. First, by anticipating future events, we may be able to take steps now to avoid engative outcomes or increase the likelihood of positive outcomes. That is, expectations have the capacity to motivate an organism. Second, even if we are unable to influence the course of future events, expectations allow us to prepare in appropriate ways. For example, we can adopt a state of arousal that is more suited to what is likely to happen next. We can also orient toward an anticipated stimulus, and so increase the speed and accuracy of future perceptions. That is, expectation allows us to prepare in

advance suitable motor responses and craft suitable perceptual strategies. Finally, expectation provides the test-bed against different representations can be evaluated.

Conclusion

The main theoretical points of this study can be summarized as follows:

1. The ability to anticipate future events is important for survival. It is reasonable to assume that evolution by natural selection has shaped perceptual and cognitive systems so that they endeavor to anticipate future events. "All brains are, in essence, anticipation machines." (Dennett, 1991; p.177).

2. It is possible to form relatively accurate expectations only because real-world environments exhibit structure and are not totally chaotic.

3. Some expectations are formed through conscious thought or reflection, as when a knowledgeable jazz listener anticipates a drum solo following a bass solo. However, most expectations are unconscious, automatic, and ubiquitous. We cannot "turn off" the mind's tendency to anticipate events, and we are usually unaware of the mind's disposition to make predictions. Except when we are surprised, or when the outcomes are important, we may not be cognizant of the specific predictions our minds make.

4. Minds are disposed to anticipate all types of stimuli -- even those stimuli (like music) which appear to be unimportant for survival.

5. Theoretically, expectations might have exclusively innate or learned origins. When an environment remains stable over millions of years, it is possible for efficient innate expectations to evolve. In hearing, innate functions are evident in such auditory reflexes as the *orienting response*. However, when an environment is highly variable, the capacity to form expectations through learning provides a better evolutionary strategy (Baldwin, 1896).

6. The auditory environments in which humans evolved appear to have been highly variable. Sounds that in one context might indicate danger, might, in another context, indicate opportunity. Given the great variety of auditory contexts in human experience, it should not be surprising that the existing research implicates learning as the preeminent source of auditory expectations.

7. Ideally, the principles underlying expectations would precisely reflect the actual principles that cause the environment to be a particular way (i.e., Shepard's *complementarity*).

8. Whether innate or learned, expectations can be formed through exposure to an environment. Expectations arise through a process of induction, in which generalizations are formed from a finite number of specific experiences.

9. Since inductive inference is known to be fallible, the generalizations formed through listener experience are also fallible. That is, the principles underlying expectations are likely to be imperfect approximations of the actual principles shaping the world (von Hippel, 2002).

10. For a broad sample of melodies, several simple principles have been identified that appear to underly the objective organization. One principle is the tendency for successive pitches to be relatively close. Experienced listeners appear to form an appropriate expectation for pitch proximity. A second principle is for pitches to exhibit a central tendency. A mathematical consequence of central tendency is the phenomenon of regression-to-the-mean. However, experienced listeners do not form an appropriate expectation for melodic regression. Instead, experienced listeners expect post-skip reversal -- which is an approximation of melodic regression. A third principle is that large intervals tend to ascend. The more common repercussion is that small intervals tend to descend. However, experienced listeners do not form the appropriate expectations. Instead, experienced listeners expect step-inertia -- which appears to arise from a combination of the tendency for pitch proximity, and the tendency for intervals to descend.

11. In a stable environment, the most frequently occurring events of the past are the most likely events to occur in the future. A simple yet optimum inductive strategy is to expect the most frequent event. The simple frequency of isolated events ("zereoth-order

distribution") forms the foundation for learned expectations.

12. An example of frequency-dependent learning in music is listener sensitivity to the distribution of scale degrees as documented by Krumhansl and elaborated by Aarden.

13. In addition to zeroeth-order frequencies, listeners are also able to learn contingent frequencies of neighboring or co-occurring events. The distance separate contingent events can range from immediate neighbors to long-range relationships. In addition, contingent probabilities can be influenced by the number of prior events that combine to influence a particular ensuing event. These probability "frames" can range from a single preceding event (first-order probability), to many preceding events (higher-order probabilities).

14. An example of contingent-frequency learning in music can be found in scale-degree successions, such as the tendency for chromatic tones to be anchored to neighboring diatonic tones.

15. Expectations provoke emotional responses. Three response categories can be distinguished: (1) responses that precede the outcome (anticipatory affective responses), (2) responses evoked by the outcome itself (secondary affective responses), and (3) responses related to the accuracy of the expectation (primary affective responses). A positively valenced primary affect ensues when an expectation proves accurate, whereas a negatively valenced primary affect ensues

when an expectation prove inaccurate.

16. Expectations that prove to be correct represent successful mental functioning. Successful anticipations help us prepare appropriate motor responses, inhibit or suppress inappropriate responses, and better perceive ensuing stimuli. Successful expectations evoke a *primary affective reward*.

17. Successful expectations can be measured. When a person's expectations are correct, they will be faster and more accurate in processing information related to the expectation. Accurate expectations can be regarded as functionally equivalent to perceptual *priming*.

18. Expectations that prove to be incorrect represent failures of mental functioning. Unsuccessful expectations evoke a *primary affective punishment* in the form of stress.

19. Stress is also evoked under situations of high uncertainty. That is, stress can ensue when we already anticipate that we will fail to anticipate events (*negative anticipatory affect*).

20. Since successful predictions evoke a positive primary affective response, we may mistakenly attribute the positive feelings to the outcome itself. That is, we may prefer a predicted outcome.

21. In addition, if we repeatedly make successful predictions for a given outcome, then the predicted outcome can itself become associated with the positive

feelings.

22. Since we are more likely to successfully predict high frequency events, it is high frequency events that tend to become associated with the primary affective reward that accompanies successful prediction. Over time, we come to prefer the high frequency events (*expectancy effect*).

23. An example of the *expectancy effect* in music is the phenomenon of tonality. Once a tonal center is established, the listener will experience the tonic stimulus as more pleasant or preferable to other states.

24. Another example of the *expectancy effect* is found in the phenomenon of meter. Once a metrical context is established, the listener will experience events that occur at the most expected moments to be more pleasant or preferable to other states.

25. Emotions can also be evoked by the outcome itself. Outcomes might be *a priori* judged as positive, negative, or neutral. It is assumed that evoked emotions tend to slowly decay in intensity following the outcome.

26. A sequence of events might evoke a mixed succession of positive and negative states. Since positively valenced states are preferred, it is advantageous for positive states to be sustained longer than negative states. Said another way, it would be advantageous for negative states to be quickly followed by a new state, whereas positive states would induce a delay before the

next state.

27. Successive events often occur in groups or segments, such as evident in phrases or entire works. In light of the above observations, listeners should prefer segments to be closed with a positive state since this increased the total positive valence.

28. If most segments are terminated with a positive state, then listeners should learn to associate positive states with closure. Closure implies repose and stability. Therefore, frequently occurring states ought to u

29. By way of summary, we can identify the following causal sequence:
    - frequently occurring events provide the best predictions for future states
    - since successful predictions are rewarded, frequently occurring events tend to become associated with positive emotions; (nominally "neutral" stimuli may thus acquire a positive valence)
    - it is preferrable for long-duration states to have a positive valence
    - by definition, the terminating event in a sequence is a long state; in creating a sequence of states, pleasure is increased if frequently occurring events tend to be placed at the ends of segments
    - through repeated exposure, terminating events become associated with closure and respose or stability; hence frequently occurring events tend to become associated with closure and repose/stability.

    In other words, frequently occurring events have a tendency to be (1) the most predicted stimulus, (2) the

most preferred stimulus, (3) the stimulus that most implies closure, and (4) the stimulus most associated with repose or stability.

30. While expected events are generally preferred, highly predictable environments can lead to reduced attention and lowered arousal -- often leading to sleepiness.

31. Apart from the simple frequency of occurrence, we are also sensitive to the co-occurrences of various events. That is, we form expectations based on conditional probabilities.

32. Most conditional probabilities reflect short-range moment-to-moment contingencies, as when one note tends to immediately follow another. However, long-range conditional probabilities may also be formed -- provided such long-range structures exist in the environment.

33. Expectations can be learned dynamically. That is, listening to a passage can help listeners form expectations that arise uniquely from the immediately preceding experience.

34. Regularities in the world are often evident only in particular contexts or environments. It is important for an organism to learn to distinguish these different environments, and to protect learned expectations within each context from the undue influence of learned associations that pertain to a different context (Cosmides & Tooby, 2000).

35. Such cognitive firewalls permit listeners to distinguish different *kinds* of musical experiences. Learned expectations can be segregated into different expectational sets or "schemas."

36. Due to lack of experience or possible cognitive deficits, it is possible that a listener fails to distinguish two forms of musical experience that other listeners experience as distinct *kinds*. A given listener might consequently experience a musical genre in a unique or idiosyncratic manner.

37. Complex stimuli may unfold in an invariant way, as when we hear the succession of pitches of *Happy Birthday.* In this case we form *veridical* expectations -- given these eight notes, the ninth note will undoubtedly be ...

38. Veridical expectations do not suppress the effects of schematic expectation (Bharucha). Schematic expectations are tenacious. This explains the apparent paradox of how some events can be both simultaneously surprising and unsurprising. For example, a wholly expected deceptive cadence doesn't entirely lose it's "deceptive" character.

39. Schemas may include prediction rules, such as the rule that successive tones tend to be close in pitch. These rules arise because they are broadly successful in their predictions (though not infallible). Some prediction rules are sub-optimum. An example is the rule for post-skip reversals. This rule is generally successful in its predictions, however the rule merely approximates a

more fundamental property of musical structure, namely that melodies tend to be constrained in their ranges. A regression-to-the-mean rule would allow listeners to better predict successive melodic pitches, however listeners appear to learn the less accurate post-skip reversal prediction rule.

40. Expectations rely on underlying mental representations. Representations might include absolute pitch, pitch-class, scale degree, interval, contour, etc. Several representations may operate concurrently in the forming of expectations. It appears that not every listener has access to all of these representations. For example, people with absolute pitch are able to code events and expectations according to absolute pitch. A major difference between people who have AP and those who don't is that AP possessors heard musical works in early life that are always in the same key, whereas non-AP possessors typically experienced musical works in a multitude of keys. It is possible, as argued by Abramson at the beginning of the twenthieth century, that the practice of singing songs in different keys, reduces the value of coding absolute pitch, and so pitch height lost its predictive value for some listeners -- leading to the ignoring of pitch height information.

41. Since more than one representation may be involved in forming expectations, an expectation may be *mixed*. For example, one element (such as pitch) may be highly unexpected, whereas another element (such as onset time) may be highly expected.

42. When the circumstances are appropriate, listeners may come to expect the unexpected. That is, a sort of "reverse psychology" may arise. Twelve-tone music has been shown to be organized in a manner consistent with such reverse psychology.

43. *Paradoxical expectations* can arise when *schematic* and *veridical* expectations differ.

44. Different listeners may have different expectations. Individual differences may be attributable to four possible sources. (1) Listeners may differ in their underlying representation codes. For example, one listener may favor an absolute pitch representation, whereas another listener favors a scale degree representation. (2) Listeners differ in the exposure to music, and so some listeners may have had less opportunity to develop appropriate schemas. (3) A listener may fail to distinguish expectational sets that may be appropriate for different genres of music. For example, as Krumhansl has shown, a listener may continue to apply a tonal schema to an atonal listening experience. (4) Listeners may differ in the accuracy of the prediction rules. For example, it is theoretically possible that a listener experiences melodic contours in accordance with the regression-to-the-mean rule rather than the post-skip-reversal rule. (5) It is theoretically possible that existing schemas may prevent a listener from distinguishing a separate schema. For example, a hypothetical scale schema `B' might interfere with the acquiring of a similar (yet distinct) schema `A'. A listener who acquires schema `A' first may retain the ability to acquire schema `B', whereas a listener who

acquires schema `B' first may be incapable of acquiring schema `A'. For example, Meyer (1956; p.46) cites the Fox Strangways who claims that some Indian music uses a scale that is very similar to the Western major scale, yet the "tonic" pitches do not coincide. The Western listener may therefore hold expectations that are wholly inappropriate to the Hindustani music (Fox Stangways, 1914; p.18).

45. The psychological responses to expectation can be classified into four categories. In the *pre-outcome* phase, an individual might imagine different possible outcomes and vicariously experience some of the feelings that would expected for each outcome. This *imaginative response* provides an important mechanism for motivating an individual to take courses of action that increase the likelihood of a positive outcome.

46. Also in the *pre-outcome* phase, appropriate arousal and attention states need to be evoked in preparation for the outcome. This *tension response* tailors the arousal and attention to match the degree of uncertainty and the importance of the possible outcomes. Obvious and inconsequential outcomes will evoke little response. Highly important yet uncertain outcomes will evoke a significant response. The response becomes more marked as the anticipated moment of the outcome approaches. The tension response is commonly manifested as stress.

47. In the *post-outcome* phase, the accuracy of an individual's predictions are appraised in the *prediction response*. A positive response will occur when the

outcome matches the individual's expectation. A negative response arises when the outcome is unexpected.

48. Finally, an emotional response will be evoked according to an appraisal of the final outcome state. A positive *outcome response* will arise if the outcome is positively appraised.

49. Primary and secondary affective responses interact. Highly predictable outcomes evoke less response than highly unpredictable outcomes. For example, an unexpected positive outcome will feel better than a highly expected positive outcome. Similarly, an unexpected negative outcome will feel worse than a highly expected negative outcome. In effect, increased uncertainty tends to amplify the aggregate affective response.

50. The delaying of an outcome has the effective of decreasing its certainty. Consequently, delay amplifies the aggregate affective response. The effect of delay is most marked when events seem to be most certain.

51. Many performance and compositional techniques can be regarded as efforts to delay expected outcomes. Such delaying techniques tend to be used in the most stereotypic musical passages.

52. The fact that learning plays a preeminent role in forming expectations, in addition to the fact that expectations can adapt dynamically to ongoing stimuli, suggests that there exist considerable opportunities to

craft a range of musics for which listeners may form appropriate expectations.

A number of questions remain to be addressed in future research concerning musical expectations. Perhaps the premiere unresolved question concerns the nature of the mental representations that underly musical expectations. *What* do listeners expect? Do they expect intervals, pitches, pitch-classes, scale degrees, scale degree successions, contours, rhythms, pitch-rhythms, etc. The existing research provides evidence that mental representations for music consist of a complex combination of musical elements. There is also evidence that different listeners may make use of different representations.

Under what circumstances are new expectational sets formed. That is, when will the auditory system erect a cognitive firewall to allow the formation of a new music-related schema? Is is possible for past listening experiences to prevent a listener from forming a new musical schema? Is it possible, for example, with the right regime of musical exposure, for a modern listener to form a truly "medieval" way of hearing early music?

Finally, what types of musical structures or principals of organization will fail to evoke appropriate learning? [8]

References

Aarden, B. (2001). *An empirical study of chord-tone doubling in common era music*. Masters thesis. School of Music, Ohio State University.

Aarden, B. (2002). Expectancy vs. retrospective perception: Reconsidering the effects of schema and continuation judgments on measures of melodic expectancy. *Proceedings of the 7th International Conference on Music Perception and Cognition*, C. Stevens, D. Burnham, G. McPherson, E. Schubert, J. Renwick (Eds.). Adelaide: Causal Productions, pp.?? -??.

Abe, J., & Oshino, E. (1990) (1990). Schema driven properties in melody cognition: Experiments on final tone extrapolation by music experts. *Psychomusicology*, Vol. 9, pp. 161-172.

Baldwin, J.M. (1896). A new factor in evolution. *American Naturalist,* Vol. 30, pp. 441-451, 536-553.

Baldwin, J.M. (1909). *Darwin and the Humanities.* Baltimore: Review Publishing.

Barnes, R. & Jones, M.R. (in press). Expectancy, attention, and time. *Cognitive Psychology.*

Berger, J. (1990). A theory of musical ambiguity. *Computers in Music Research,* Vol. 2, pp. 91-119.

Bharucha, J. (1984). Anchoring effects in music: The resolution of dissonance. *Cognitive Psychology,* Vol. 16, pp. 485-518.

Bharucha, J. (1987). MUSACT: A connectionist model of musical harmony. In *Proceedings of the Ninth Annual Conference of the Cognitive Science Society*, Hillsdale, NJ: Erlbaum, pp. 508-517.

Bharucha, J. (1994). Tonality and expectation. In R. Aiello (Ed.), *Musical Perceptions,* Oxford: Oxford University Press, pp. 213-239.

Bharucha, J. (1996). Melodic anchoring. *Music Perception,* Vol. 13, No. 3, pp. 383-400.

Bharucha, J.J., & Stoeckig, Keiko. (1987). Priming of chords: Spreading activation or overlapping frequency spectra? *Perception & Psychophysics,* Vol. 41, No. 6, pp. 519-524.

Bigand, E., & Pineau, M. (1997). Global context effects on musical expectancy. *Perception & Psychophysics,* Vol. 59, No. 7, pp. 1098-1107.

Bornstein, R.F. (1989). Exposure and affect: Overview and meta-analysis of research, 1968-1987. *Psychological Bulletin,* Vol. 106, No. 2, pp. 265-289.

Carlsen, J.C. (1981). Some factors which influence melodic expectancy. *Psychomusicology,* Vol. 1, pp. 12-29.

Chernoff, J.M. (1979). *African Rhythm and African Sensibility: Aesthetics and Social Action in African Musical Idioms.* Chicago: University of Chicago Press.

Cohen, A.J. (1982). Exploring the sensitivity to structure in music. *Canadian University Music Review,* Vol. 3, pp. 15-30.

Cohen, J.E. (1962). Information theory and music. *Behavioral Science,* Vol. 7, No. 2, pp. 137-163.

Cook, N. (1987). The perception of large-scale tonal closure. *Music Perception,* Vol. 5, No. 2, pp. 197-206.

Coons, E., & Kraehenbuehl, D. (1958). Information as a measure of structure in music. *Journal of Music Theory,* Vol. 2, pp. 127-161.

Cosmides, L., & Tooby, J. (2000). Consider the source: The evolution of adaptations for decoupling and metarepresentations. In: D. Sperber (Ed.) *Metarepresentations: A Multidisciplinary Perspective.* Oxford: Oxford University Press, pp. 53-115.

Cuddy, L.L., & Lunney, C.A. (1995). Expectancies generated by melodic intervals: Perceptual judgments of melodic continuity. *Perception & Psychophysics,* Vol. 57, No. 4 pp. 451-462.

Damasio, A. (1994). *Descartes' Error: Emotion, Reason, and the Human Brain.* G.P. Putnam's Sons.

Dennett, D.C. (1991). *Consciousness Explained.* Boston: Little, Brown.

Deutsch, D. (1978). Delayed pitch comparisons and the principle of proximity. *Perception & Psychophysics,* Vol. 23, pp. 227-230.

Dowling, W.J. (1967). *Rhythmic fission and the perceptual organization of tone sequences.*

Unpublished doctoral dissertation, Harvard University, Cambridge, MA.

Dowling, W.J. (1978). Scale and contour: Two components of a theory of memory for melodies. *Psychological Review.* Vol. 85, pp. 341-354.

Dowling, W.J., & Harwood, D.L. (1986). *Music Cognition.* San Diego: Academic Press.

Dusman, L. (1994). Unheard-of: Music as performance and the reception of the new. *Perspectives of New Music*, Vol. 32, No. 2, pp. 130-146.

Federman, F. (1996). *A study of various representations using NEXTPITCH: A learning classifier system.* PhD dissertation, City University of New York.

Feldmann, C.M. (1997). *Erwartungsdiskrepanz und emotionales Erleben von Musik.* PhD dissertation, Ruprecht-Karls-U., Heidelberg. Hildesheim: Olms. ISBN: 3-487-10710-4.

Fox Stangways, A.H. (1914). *The Music of Hindostan.* London: Oxford University Press.

Francès, R. (1988). *La perception de la musique.* (1958) Translated by W.J. Dowling as *The Perception of Music* Hillsdale, N.J.: Lawrence Erlbaum Associates.

Gallistel, C.R. (1990). *The Organization of Learning.* Cambridge, MA: MIT Press.

Gardner, M. (1978). Mathematical Games -- white and brown music, fractal curves and one-over-f fluctuations. *Scientific American,* Vol. 238, No. 4, pp. 16-32.

Gjerdingen, R.O. (1988). *A Classic Turn of Phrase: Music and the Psychology of Convention.* Philadelphia: University of Pennsylvania Press.

Goldstone, J.A. (1979). *A general mathematical theory of expectation models of music.* PhD dissertation, University of Southern California.

Gotlief, H., & Konecni, V.J. (1985). The effects of instrumentation, playing style, and structure in the Goldberg variations by Johann Sebastian Bach. *Music Perception,* Vol. 3, No. 1, pp. 87-102.

Granot, R., & Donchin, E. (2002). Do re mi fa sol la ti - Constraints, congruity, and musical training: An event-related brain potentials study of musical expectancies. *Music Perception*, Vol. 19, No. 4, pp. 487-528.

Greenberg, D., & MacMillan, S. (1996). *Bach Meets Cape Breton.* [Audio CD]. Canada: Marquis Records #181. ASIN: B000003WHK.

Hansen, F. (1969). Musical meaning as pointing. *Music Review,* Vol. 30, No. 4, pp. 300-307.

Hasher, L., & Zacks, R.T. (1984). Automatic processing of fundamental information. *American Psychologist,* Vol. 39, pp. 1372-1388.

Hick, W.E. (1952). On the rate of gain of information. *Quarterly Journal of Experimental Psychology,* Vol. 4, pp. 11-26

von Hippel, P. (1998). *Post-skip reversals reconsidered: Melodic practice and melodic psychology.* PhD dissertation, Stanford University.

von Hippel, P. (2000). Redefining pitch proximity: Tessitura and mobility as constraints on melodic intervals. *Music Perception,* Vol. 17, No. 3, pp. 315-327.

von Hippel, P. (2000). Questioning a melodic archetype: Do listeners use gap-fill to classify melodies? *Music Perception,* Vol. 18, No. 2, pp. 139-153.

von Hippel, P., & Huron, D. (2000). Why do skips precede reversals? The effect of tessitura on melodic structure. *Music Perception,* Vol. 18, No. 1, pp. 59-85.

Huron, D. (2001a). What is a musical feature? Forte's analysis of Brahms's Opus 51, No. 1, Revisted. *Music Theory Online,* Vol. 7, No. 4. **Online text.**

Huron, D. (2001b). Tone and Voice: A derivation of the rules of voice-leading from perceptual principles. *Music Perception,* Vol. 19, No. 1, pp. 1-64. **Online text.**

Huron, D., & von Hippel, P. (2000). Tonal and Contra-tonal structure of Viennese twelve-tone rows.

Paper presented at the Society for Music Theory Conference, Toronto, Canada.

Hyman, R. (1953). Stimulus information as a determinant of reaction time. *Journal of Experimental Psychology,* Vol. 45, pp. 423-432.

Jairazbhoy, N. (1971). *The Rags of North Indian Music: Their Structure and Evolution.* London: Faber & Faber.

Jones, M.R. (1981). Music as a stimulus for psychological motion: Part 1. Some determinants of expectancies. *Psychomusicology,* Vol. 1, pp. 34-51.

Jones, M.R. (1982). Music as a stimulus for psychological motion: Part 2. An expectancy model *Psychomusicology,* Vol. 2, pp. 1-13.

Jones, M.R. (1990). Learning and the development of expectancies: An interactionist approach. *Psychomusicology,* Vol. 9, No. 2, pp. 193-228.

Jones, M.R. (1992). Attending to musical events. In: M.R. Jones & S. Holleran (eds.) *Cognitive Bases of Musical Communication.* Washington, DC: American Psychological Association; pp. 91-110.

Jones, M.R., & Boltz, M. (1989). Dynamic attending and responses to time. *Psychological Review,* Vol. 96, pp. 459-491.

Jones, M.R., Boltz, M., & Kidd, G. (1982). Controlled attending as a function of melodic and temporal

context. *Perception & Psychophysics,* Vol. 32, pp. 211-218.

Jones, M.R., Moynihan, H., MacKenzie, N., & Puente, J. (in press). Temporal aspects of stimulus-driven attending in dynamic arrays. *Psychological Science.*

Kelly, M.H., & Martin, S. (1994). Domain-general abilities applied to domain-specific tasks: Sensitivity to probabilities in perception, cognition, and language. *Lingua,* Vol. 92, pp. 105-140.

Karno, M., & Konecni, V.J. (1992). The effects of structural interventions in the first movement of Mozart's symphony in G minor K.550 on aesthetic preference. *Music Perception,* Vol. 10, No. 1, pp. 63-72.

Kraehenbuehl, D., & Coons, E. (1959). Information as a measure of the experience of music. *Journal of Aesthetics & Art Criticism,* Vol. 17, pp. 510-522.

Kramer, J.D. (1982). Beginnings and endings in Western art music. *Canadian University Music Review/Revue de musique des universites canadiennes,* Vol. 3, ????

Krumhansl, C. (1979). The psychological representation of musical pitch in a tonal context. *Cognitive Psychology,* Vol. 11, pp. 346-374.

Krumhansl, C. (1990). *Cognitive Foundations of Musical Pitch.* Oxford: Oxford University Press.

Krumhansl, C. (1995). Effects of musical context on similarity and expectancy. *Systematische Musikwissenschaft/Systematic Musicology/Musicologie systematique*, Vol. 3, No. 2, pp. 211-250.

Krumhansl, C. (1999). Effects of perceptual organization and musical form on melodic expectancies. In M. Leman (Ed.), *Music, Gestalt, and Computing: Studies in Cognitive and Systematic Musicology.* Berlin: Springer Verlag, pp. 294-320.

Krumhansl, C., & Kessler, E.J. (1982). Tracing the dynamic changes in perceived tonal organization in a spatial representation of musical keys. *Psychological Review*, Vol. 89, pp. 334-368.

Krumhansl, C., Sandell, G.J., & Sergeant, D.C. (1987). The perception of tone hierarchies and mirror forms in twelve-tone serial music. *Music Perception*, Vol. 5, pp. 153-184.

Kunst-Wilson, W.R., & Zajonc, R.B. (1980). Affective discrimination of stimuli that cannot be recognized. *Science,* Vol. 207, pp. 557-558.

Lang, P.J., Simons, R.F., & Balaban, M.T. (Eds.). (1997). *Attention and Orienting: Sensory and Motivational Processes.* Mahwah, New Jersey: Lawrence Erlaum Associates.

Larson, S. (1999). Continuations as completions: Studying melodic expectation in the creative microdomain Seek Well. In M. Leman (Ed.), *Music,*

Gestalt, and Computing: Studies in Cognitive and Systematic Musicology. Berlin: Springer Verlag, pp. 321-334.

Lerdahl, F., & Jackendoff, R. (1983). *A Generative Theory of Tonal Music*, Cambridge, MA: MIT Press.

Lewis, M. (1995). Self-conscious emotions. *American Scientist*, Vol. 83, No. 1, pp. 68-78.

Lieberman, P. (1967) *Intonation, Perception and Language.* Cambridge, MA: MIT Press.

Lord, A. (1960). *The Singer of Tales*. Cambridge, MA: Harvard University Press.

Mandler, G. (1975). *Mind and Emotions.* New York: J. Wiley.

Manzara, L.C., Witten, I.H., & James, M. (1992). On the entropy of music: an experiment with Bach chorale melodies. *Leonard Music Journal* Vol. 2, No. 1, pp. 81-88.

McCredie, A.D. (1983). Some concepts, constructs, and techniques in comparative literature and their interface with musicology. *International Review of the Aesthetics and Sociology of Music,* Vol. 14, No. 2, pp. 147-165.

Merriam, A.P., Whinery, S., & Fred, B.G. (1956). Songs of a Rada community in Trinidad. *Anthropos* 51, 157-174.

Meyer, L.B. (1956). *Emotion and Meaning in Music.* Chicago: University of Chicago Press.

Miyazaki, K. (1990). The speed of musical pitch identification by absolute-pitch possessors. *Music Perception,* Vol. 8, No. 2, pp. 177-188.

Moles, A. (1958/1966). *Théorie de l'information et perception esthétique* Paris, 1958. Trans. as *Information Theory and Aesthetic Perception.* Urbana, IL: University of Illinois Press, 1966.

Moreland, R.L., & Zajonc, R.B. (1977). Is stimulus recognition a necessary condition for the occurrence of exposure effects? *Journal of Personality and Social Psychology,* 35, 191-199.

Moreland, R.L., & Zajonc, R.B. (1979). Exposure effects may not depend on stimulus recognition. *Journal of Personality and Social Psychology,* 37, 1085-1089.

Narmour, E. (1990). *The Analysis and Cognition of Basic Melodic Structures: The Implication-Realization Model.* Chicago: University of Chicago Press.

Narmour, E. (1991). The influence of embodied registral motion on the perception of higher-level melodic implication. In M.R. Jones & S. Holleran (Eds.), *Cognitive Bases of Musical Communication.* Washington, DC: American Psychological Association, pp. 69-90.

Narmour, E. (1992). *The Analysis and Cognition of Melodic Complexity: The Implication-Realization*

Model. Chicago: University of Chicago Press.

Narmour, E. (1999). Hierarchical expectation and musical style. In D. Deutsch (Ed.), *The Psychology of Music.* 2nd edition. San Diego: Academic Press, pp. 441-472.

Narmour, E. (2000). Music expectation by cognitive rule-mapping. *Music Perception,* Vol. 17, No. 3, pp. 329-398.

Olson, J.M., Roese, N.J., & Zanna, M.P. (1996). Expectancies. In: Higgens, E.T. & Kruglanski, W. (eds.) *Social Psychology: Handbook of Basic Principles.* New York: Guilford Press, pp. 211-238.

Ortmann, O.R. (1926). *On the melodic relativity of tones.* Princeton, NJ: Psychological Review Company. (Vol. 35, No. 1 of *Psychological Monographs.)*

Palmer, C., & Krumhansl, C. (1990). Mental representations for musical meter. *Journal of Experimental Psychology: Human Perception and Performance*, Vol. 16, No. 4, pp. 728-741.

Parry, M. (1971). *The Making of Homeric Verse: The Collected Papers of Milman Parry*. Oxford: Clarendon Press.

Perrott, D., & Gjerdingen, R.O. (1999). Scanning the dial: An exploration of factors in the identification of musical style. Paper presented at the Society for Music Perception and Cognition Conference, Evanston, IL.

Pike, K.L. (1945) *The Intonation of American English.* Ann Arbor, Michigan: University of Michigan Press.

Pinkerton, R.C. (1956) Information theory and melody. *Scientific American*, Vol. 194, No. 2, pp. 77-86.

Reber, A.S. (1993). *Implicit Learning and Tacit Knowledge: An Essay on the Cognitive Unconscious.* Oxford: Oxford University Press.

Regnault, P., Bigand, E., & Besson, M. (1999). Sensory and cognitive influences on musical expectancy revealed by event-related potentials. *Psychophysiology*, Vol. 36, S17-S17, Supplement 1.

Reimer, B. (1964). Information theory and the analysis of musical meaning. *Council for Research in Music Education*, Vol. 2, pp. 14-22.

Rosch, E. (1973). On the internal structure of perceptual and semantic categories. In T. Moore (ed.), *Cognitive development and the acquisition of language.* New York: Academic Press.

Rosch, E. (1975). Cognitive reference points. *Cognitive Psychology,* Vol. 7, pp. 532-547.

Rosner, B.S., & Meyer, L.B. (1982). Melodic processes and the perception of music. In D. Deutsch (Ed.), *The Psychology of Music.* New York: Academic Press, pp. 317-341.

Saffran, J.R., Newport, E.L., & Aslin, R.N. (1996). Word segmentation: the role of distributional cues. *Journal of Memory and Language,* Vol. 35, pp. 606-621.

Saffran, J.R., Johnson, E.K., Aslin, R.N., & Newport, E.L. (1999). Statistical learning of tone sequences by human infants and adults. *Cognition,* Vol. 70, pp. 27-52.

Schellenberg, E.G. (1996). Expectancy in melody: Tests of the implication-realization model. *Cognition,* Vol. 58, pp. 75-125.

Schellenberg, E.G. (1997). Simplifying the implication-realization model. *Music Perception,* Vol. 14, No. 3, pp. 295-318.

Schmuckler, M.A. (1988). *Expectation in music: Additivity of melodic and harmonic processes*. PhD dissertation, Cornell University.

Schmuckler, M.A. (1989). Expectation in music: Investigation of melodic and harmonic processes. *Music Perception,* Vol. 7, pp. 109-150.

Schmuckler, M.A. (1990). The performance of global expectation. *Psychomusicology,* Vol. 9, pp. 122-147.

Shannon, C.E. (1948). A mathematical theory of communication. *Bell System Technical Journal*, Vol. 27, pp. 379-423, 623-656.

Shannon, C.E., & Weaver, W. (1949). *The Mathematical Theory of Communication.*, Urbana,

IL: University of Illinois Press.

Shepard, R.N. (1981). Psychophysical complementarity. In: M. Kubovy & J.R. Pomerantz (eds.), *Perceptual Organization*., Hillsdale, NJ: Erlbaum, pp. 279-341.

Simpson, J. (1996). A formal analysis of note-interdependence in selected works. Unpublished manuscript. Available **online text**.

Simpson, J., & Huron, D. (1994). Absolute pitch as a learned phenomenon: Evidence consistent with the Hick-Hyman law. *Music Perception,* Vol. 12, No. 2, pp. 267-270.

Sloboda, J.A. (1991). Music structure and emotional response: Some empirical findings. *Psychology of Music,* Vol. 19, No. 2, pp. 110-120.

Sloboda, J.A. (1992). Empirical studies of emotional response to music. In: M.R. Jones & S. Holleran (eds.), *Cognitive Bases of Musical Communication*. Washington, DC: American Psychological Association, pp.33-50.

Takeuchi, A.H., & Hulse, S.H. (1991). Absolute pitch judgments of black- and white-key pitches. *Music Perception*, Vol. 9, pp. 27-46.

Tekman, H.G. (1998). Effects of melodic accents on perception of intensity. *Music Perception*, Vol. 15, No. 4, pp. 391-401.

Tekman, H.G. (2001). Accenting and detection of timing variations in tone sequences: Different kinds of accents have different effects. *Perception & Psychophysics*, Vol. 63, No. 3, pp. 514-523.

't Hart, J., Collier, R., & Cohen, A. (1990). *A Perceptual Study of Intonation; An Experimental-phonetic Approach to Speech Melody.* Cambridge: Cambridge University Press.

Thompson, W.F. Balkwill, L.L., & Vernescu, R. (2000). Expectancies generated by recent exposure to music. *Memory and Cognition*, Vol. 28, No. 4, pp. 547-555.

Thompson, W.F. Cuddy, L.L., & Plaus, C. (1997). Expectancies generated by melodic intervals: Evaluation of principles of melodic implication in a melody production task. *Perception & Psychophysics*, Vol. 59, No. 7, pp. 1069-1076.

Thompson, W.F., & Stainton, M. (1998). Expectancy in Bohemian folk song melodies: Evaluation of implicative principles for implicative and closural intervals. *Music Perception*, Vol. 15, No. 3, pp. 231-252.

Tillmann, B., Bigand, E., & Pineau, M. (1998). Effects of global and local contexts on harmonic expectancy. *Music Perception*, Vol. 16, No. 1, pp. 99-117.

Toiviainen, P., Krumhansl, C., Luohivuori, J., Jarvinen, T., & Eerola, T. (1999). Melodic expectation in Finnish spiritual folk hymns: Convergence of

statistical, behavioral, and computational approaches. *Music Perception*, Vol. 17, No. 2, pp. 151-195.

Unyk, A.M. (1990). An information-processing analysis of expectancy in music cognition. *Psychomusicology,* Vol. 9, No. 2, pp. 229-240.

Unyk, A.M., & Carlsen, J.C. (1987). The influence of expectancy on melodic perception. *Psychomusicology,* Vol. 7, pp. 3-23.

Vos, P.G., & Troost, J.M. (1989). Ascending and descending melodic intervals: Statistical findings and their perceptual relevance. *Music Perception,* Vol. 6, No. 4, pp. 383-396.

Voss R.F., & Clarke J. (1975). 1/f noise in music and speech. *Nature,* 258, pp. 317-318.

Voss R.F., & Clarke J. (1978). 1/f noise in Music: Music from 1/f Noise. *Journal of the Acoustical Society of America*, Vol. 63, No. 1, pp. 258-263.

Watt, H.J. (1924). Functions of the size of interval in the songs of Schubert and of the Chippewa and Teton Sioux Indians. *British Journal of Psychology,* Vol. 14, pp. 370-386.

Werbik, H. (1969). L'indétermination et les qualités impressives des modèles stimulants mélodiques. [Uncertainty and the affective qualities of melodic stimuli.] *Sciences de l'Art,* 1-2, pp. 25-37.

Wilson, W.R. (1975). Unobtrusive induction of positive attitudes. Ph.D. dissertation, University of Michigan.

Wilson, W.R. (1979). Feeling more than we can know: Exposure effects without learning. *Journal of Personality and Social Psychology,* Vol. 37, pp. 81-821.

Wittgenstein, L. (1966). *Lectures and Conversations on Aesthetics, Psychology and Religious Belief.* Compiled from notes taken by Yorick Symthies, Rush Rhees, and James Taylor; Cyril Barret (Ed.). Oxford: Blackwell.

Wong, A.K.C., & Ghahraman, D. (1975). A statistical analysis of interdependence in character sequences. *Information Sciences,* Vol. 8, pp. 173-188.

Youngblood, J.E. (1958). Style as information. *Journal of Music Theory,* Vol. 2, pp. 24-35.

Schemas

Listening is not a passive activity where we simply classify successive stimuli as we encounter them. Listening is an active process. When we listen to a spoken sentence, for example, we formulate hypotheses about what is being said. We anticipate what will happen next. The context of an utterance prepares us for possible outcomes. We may already have an idea of what someone will say before they begin to speak.

In the case of music, we need only a few seconds of exposure to situate a musical work according to genre, tempo, meter, and so forth. Within three or four seconds, we will know whether the music is fast or slow, whether the key is major or minor, and whether it is baroque, bebop, big-band or blue-grass. Within a few more seconds, we will have a good intuition of the *scenario* of the music -- what is likely to happen, how the work may end, etc. Such scenarios act like "templates" that help us to orient ourselves during the listening experience.

Through years of listening experiences, each listener develops a repertoire of such possible scenarios. Such mental preconceptions of the normal course of events are referred to as *schemas*.* More precisely, a schema may be defined as a knowledge structure that arises from past experience, and which influences how we perceive and interpret current events. In a sense, schemas are like archetypal "stories" -- such as love stories, tragedies, horror, comedies, etc. Whether or not we are consciously aware of it, we will have intuitions about what is likely to happen in these stories. For example, love stories always have some impediment that must be overcome in order for the lovers to get together. Both action films and comedy films tend to have a chase scene near the end.

Schemas don't simply apply to the overall patterns in musical works. Individual phrases, and even note-to-note successions tend to follow certain norms. James Carlsen and his colleagues have carried out a number of experiments mapping-out what listeners expect to happen next -- given various antecedent musical events.

Expectation

When experiencing music, listeners are not merely passive observers. At an unconscious level, listeners form expectations about what will happen next. Some of these expectations are obvious to a listener. That is, the fact that we are expecting something rises into consciousness, and we are aware that we are expecting something. These expectations are based on our past experiences. -innate, learned, veridical, schematic, enculturated, Not all learned schematic expectations are "cultural." Some will arise from idiosyncratic personal listening habits. For example, a lover of Bebop jazz may form some Bebop-specific expectancies. Yet our Bebop lover may have little or no interaction with other Bebop fans, and so the social or group component so commonly regarded as the touch-stone of "culture" may be absent. Other learned expectations arise from stimuli that are commonplace throughout the world and so learned-by-transcultural. For example, with the exception of Swiss yodelling, "melodies" throughout the world have a strong tendency for small pitch motions ("pitch proximity"). Therefore, an expectation for small pitch intervals cannot be considered "cultural," even though it is learned. Expectations are evident -Meyer (1956) -tendency tones (play a scale up to `ti') -expectation -expectation dissonance -schematic and veridical expectations

Veridical Expectations

An expectation that arises due to knowledge about a specific stimulus, such as familiarity with a given musical work. When a listener expects a certain note in a well-known song, the expectation may be regarded as veridical.

By contrast, when a listener exhibits a general expectation for the leading-tone to be followed by the tonic, the expectation is regarded as a schematic expectation.

Schematic Expectations

An expectation that arises due to the existence of a mental schema. When a listener has a general expectation for the leading-tone to be followed by the tonic, the expectation may be regarded as schematic. Contrast with veridical expectation. Paul von Hippel carried out a detailed experiment on melodic expectation For large melodic intervals, musician listeners expect a change of direction. A large ascending leap, for example, causes an expectation for an ensuing lower pitch. For small melodic intervals (1 or 2 semitones), there is an expectation for the melody to continue in the same direction. The following graph illustrates these expectations for musician listeners.

Six general questions are addressed: (1) What is the biological purpose of forming expectations? (2) What aspects of musical organization do listeners anticipate? (3) How are expectations formed? (4) Do all listeners form the same expections, and if not, what accounts for the differences? (5) Are listeners' expectations accurate? and (6) How do expectations evoke emotional responses for music listeners?

It is argued that the common perceptual and emotional phenomena associated with expectation originate in the evolution of the auditory system. Many musical works are likely organized so as to evoke emotional responses that, in part, arise due to expectation-related manipulations.

Forming accurate expectations about the world is important for an organism. Like other animals, humans learn from our environments, and we form expectations of future events based on our exposure to past events. Insofar as possible, these expectations should accurately reflect reality.

When listening to music, listeners form expectations about possible future events. Our expectations are learned, but the propensity to form such expectations is innate and unconscious. As listeners, we will form musical expectations whether we want to or not. Even people who, due to injury, have lost their long-term memory, continue to learn to form new expectations on the basis of exposure.

In discussing musical expectations, we need to address a number of questions. Five questions are especially central. We have already addressed the question of why expectations exist in the first place. Second, what features of the music do listeners learn to anticipate? Third, do all listeners form the same expections, and if not, what accounts for the differences? Fourth, are our expectations accurate? What happens when, like the pacific bull-frog, our expectations prove faulty? And finally, what are the psychological consequences of forming accurate or inaccurate anticipations? Specifically, how might the dance of expectations lead to different emotions?

In recent years, psychological research has illuminated a number of aspects of musical expectations. Most of the research has focussed on melody and melodic expectations, but we'll also address harmony and rhythm later in our discussion.

Schematic versus Veridical Expectations

Of course familiarity with a single piece changes the experience of listening to the work itself. Clearly, a listener has nearly "perfect" expectations for highly familiar pieces, such as *Happy Birthday*. Cognitive psychologists distinguish two types of memory (and expectations): *veridical* and *schematic*. A veridical memory is a memory for a passage associated with a specific work. For example, the G-G-G-Eb motive is unique to Beethoven's Symphony No. 5. A schematic memory is a memory for a commonplace passage. For example, the pitch sequence *do-ti-do* occurs in a large number of works.

To illustrate the difference between veridical expectations and schematic expectations consider the following English phrases:

1. *Four score and seven years ago ...*
2. *Once upon a time ...*

The first passage is quoted from Lincoln's "Gettysburg Address" and is unique to the passage. The second passage is just as well-known, but is not unique to a particular story or fable. Most people are aware that a number of passages begin "Once upon a time" whereas there is only one continuation for "Four score and seven years ago".

When a work is perfectly known to some listener, what does it mean to have expectations? A classic problem is how a deceptive cadence can continue to sound "deceptive" when familiarity with a work makes the progression inevitable?

Having distinguished schematic versus veridical expectations, let me now withdraw and refine this distinction. There is nothing to suggest that veridical and schematic expectations are fundamentally different. A better way to think about veridical expectations is that they simply describe Markov chains containing long sequences where the note-to-note transitional probabilities equal 1.0 (or nearly so). In other words, given a specific sequence of N notes the listener's past exposure suggests a probability of nearly 1.0 for some given continuation. In short, what we have called veridical expectations are simply comparatively long stable sequences, whereas schematic expectations are shorter sequences that might have two or three plausible continuations.

One piece of evidence in support of this claim can be found in the sorts of memory errors often seen when amateur musicians play recitals or auditions. Many musical works have long sections that are repeated. A nervous performer sometimes lapses into a memory loop where they play the same passage verbatim without taking a "second ending" or otherwise continuing as they should with the rest of the piece. In effect, the music contains a long Markov chain with transitional probabilities of 1.0. However, there are boundary points where the music provides two or three choices of what should happen next. The nervous performer appears unable to break out of the chain -- seemingly perpetually doomed to take the highest probability path. Said another way, the performer's representation for the work is not truly veridical: the music is not represented as a single linear sequence of events from beginning to end. Rather, there are periodic points

where the conditional probabilities are significantly less than one, and some cognitive choice must be made.

The only difference between a veridical coding and a schematic coding is the size of the coded segments, and the fact that schematic transitions are less determinate for veridical expectations. It should not at all be surprising that many long sequences of states are unique to given musical works. Given the explosion of possible combinations for a modest number of successive events, it does not take many notes to uniquely identify one particular piece.

The point of this discussion is to note that while listeners have memories for the sequences of events that constitute an entire musical work, these memories are not qualitatively different from the memories we have for typical baroque figures, common jazz riff elements, or stereotypic country & western harmonies. The work of Parry (1971) and Lord (1960) concerning the centonization of ballads and legends similarly suggests that the way we construe "a work" may still leave considerable statistical latitude for the choice of particular segments as the work is "filled in" during performance. Finally, introspection tells us that our memory for many musical works really amounts to a handful of memorable passages. When we attempt to hum all the way through, say, Dvorak's *New World Symphony,* we find ourselves skipping large segments, or repeating ourselves in the same manner as the nervous recitalist.

The Passing Tone

	Outcome	Predictive	Tension
pre-	consonant -		moderate to low tension

passing-tone			
passing-tone	dissonant	moderate predictive success due to proximity	somewhat low tension; might return (neighbor tone) or continue in same direction
resolving	consonant	high predictive success	-

Unembellished

	Outcome	Predictive	Tension
pre-resolution	consonant	-	moderate to low tension; relatively strong expectation of the ensuing resolving pitch
resolution	consonant	moderate predictive success	-

More Material

Research has established that the primary and secondary affective responses interact with each other. These interactions are illustrated in Table 5 which provides a taxonomy of limbic/emotional responses commonly evoked by different circumstances. The secondary affect (or "outcome-related affect") is appraised as having either a positive, negative, or neutral *valence*. Positive outcomes are associated with opportunity and pleasure; negative outcomes are associated with threat and displeasure. The primary affect (or "expectancy-accuracy affect") is either expected or unexpected. (Later we will discuss the consequences of delay.)

The interactions between primary and secondary affect has been measured by Barbara Mellers and her colleagues. A simple experimental design, for example, asks amateur basketball players to take shots from different positions around the court. Before each shot, the player is asked

what they think is the likelihood of scoring the basket. Following each shot, the player is asked how good they feel. As you might expect, players are happiest when they make a shot and are unhappy when they miss a shot. However, the degree of satisfaction/dissatisfaction is directly related to the player's expectation. The greatest unhappiness occurs when the player misses a shot that they judge to be "easy" and are happiest when they score a basket that is judged to have a low probability of success. In general, unexpected fortune or misfortune cause the greatest emotional responses. That is, low expectation amplifies the emotional response to the outcome.

This relationship can be expressed through the equation given below. The value ***psi*** represents the realized subjective value when experiencing some specified outcome. This subjective value is determined by two summed terms -- the first representing the primary (expectation-related) affect, and the second representing the secondary (outcome-related) affect. The value ***v(O)*** designates the prior subjective preference for outcome ***O*** and ranges from negative values (negative valence) through positive values (positive valence); $\boldsymbol{p_e(O)}$ designates the subjective likelihood of outcome ***O***.

In the primary affect term, the subjective likelihood for outcome ***O*** is scaled so that maximum weight is given when occurrence is certain (*p()*=1.0) or when non-occurrence is certain (*p()*=0.0). A constant *k* provides a weighting for the relative importance of primary and secondary affect terms.

Emotions are evoked by a combination of what we expect will happen, what actually happens, and the accuracy of our expectations. More precisely, emotions are evoked by a combination of how we appraise the value of the expected and actual states, and how we appraise our predictive accuracy. In their model of expectation, Olson, Roese and Zanna (1996) make a useful distinction between primary and secondary affect related to expectation. Since the preeminent goal of forming expectations is to provide accurate predictions, a positive primary affect is evoked when the expectation proves accurate and a negative primary affect is evoked when the expectation proves inaccurate. Confirmation of expected outcomes generally induces a positive emotional response (Mandler, 1975). Of course it is possible to expect bad outcomes. Following a snow storm, for example, I might predict that I will slip and fall on the sidewalk. In the event that I actually fall, the outcome will feel unpleasant, but the experience will be mixed with a certain satisfaction at having correctly anticipated the outcome. It is as though brains know not to shoot the messenger: accurate expectations are to be valued (and rewarded) even when the news is bad.

Of course, outcomes are also important, and so a second affective response will result from an appraisal of the ultimate state of things. Outcomes can be appraised from a number of diffferent perspectives. Huron (2002) has distinguished six systems ranging from valenced reflexes to social appraisals. In the case of music, an outcome might evoke positive or negative responses due to differences in sensory dissonance on the one hand, or according to judgments of the social group associated with a particular style. An extensive literature exists regarding emotional

responses to particular states (REFS). It is not the purpose of this article to review this literature. We will simply assume that outcome-related emotions exist.

In addition, negative and positive outcomes can be amplified or attenuated depending on the subjective certainty of the outcome. We will note that *delay* plays an important role in amplifying the emotional valence of highly expected outcomes.

Footnotes

[1] The tonic is the most common pitch only for tonal music that does not contain modulations. Return to text.

[2] Information Theory showed early promise in the analysis of music, but was abandoned by the mid 1960s. Three factors probably contributed its demise in musical circles. In the first instance, scholars tended to rely on measures of "self-information" -- that is, probabilities that were based on the work itself. However, the theory strongly suggested that the correct way to analyse works was by using probabilities that reflect the entire musical experience of a general listener. A proper analysis would involve comparing a musical work to a large sample of other works. At the time, no large-scale musical databases existed that could be used for such analyses. In the second instance, the computers that were available to music scholars in the late 1950s and early 1960s were very slow and had limited memory capacity. Even if large musical

databases had existed, it would have proved difficult to carry out the types of analyses suggested by information theorists. In the arts and humanities, information theory was applied notably to the analysis of language text. However, in 1956, Noam Chomsky's landmark book, *Syntactic Structures*, appeared. Chomsky argued that information theory was incapable of capturing important elements of langauge organization, and offered an alternative analytic approach. The close similarity between Chomsky's *tranformational generative grammars* and Schenkerian analysis, led to a wholesale shift toward Schenkerian studies. It was not until the 1970s that it became recognized that Chomsky's criticisms of information theory were unfounded. By that point, music theorists regarded information theory as old-fashioned and irrelevant. Over the ensuing decades, information theory has continued to be an active area of research in mathematics, computer science, and communications engineering. With extensions such as m-dependency theory, information theory has grown into a remarkably powerful paradigm for analysing abstract structures, such as those found in music. Return to text.

[3] John Chernoff (1979, p. 94) provides a lovely description of how rhythmic organization pervades the west African culture of Ghana. In a customs office, Chernoff had to wait while a clerk typed copies of invoices. "Using the capitalization shift key with his little fingers to pop in accents between words, [the clerk] beat out fantastic rhythms. Even when he looked at the rough copies to find his next

sentence, he continued his rhythms on the shift key. He finished up each form with a splendid flourish on the date and port of entry. ... I realized that I was in a good country to study drumming." Return to text.

[4] A survey of European folksongs indicates that melodies in major keys are roughly twice as common as melodies in minor keys. This suggests that even the choice of initial schema may be sensitive to the frequency of occurrence of various contexts. Return to text.

[5] The term "secondary affect" is used to designate what we have called here the "outcome response". Return to text.

[6] From Robert Frost's *The Birds Do Thus*. Return to text.

[7] Recall that the imaginative response relates to the activity of contemplating different future states. There is a case to be made that the vast majority of listening is done *teleologically*. That is, music listening is dominated by a sense of inevitability, where the listener is unable to entertain alternative ways in which the music might unfold. When I became active as a composer, I was amazed that it was possible to listen to well-known compositions with a composer's sense of "choices". That is, one could listen to a composer like Beethoven with a sense that certain choices could have been different: Beethoven might have added another variation of the current theme, or brought back and theme used earlier, or that a section of the development might

have been shortened. In other words, the experience of composing made it possible to listen without the sense that the music is inevitably the way it is, and not some other way.

In ordinary music listening, it is very likely that the imaginative response is largely absent or muted. This is obviously a convenient point-of-view, since attempting to analyze the non-teleological or imaginative component to listening would be extremely daunting. Return to text.

[8] I am indebted to Paul von Hippel, Bret Aarden, Simon Durrant, Jonathan Berger, and Joy Ollen for comments made on earlier drafts of this article.
Return to text.

- **Music 829 Main Page**
- **OSU Music Cognition Home Page**

_ frequency frequency associated with closure stimuli become indicative of closure closure stability stimuli associated with closure perceived as more stable pleasure tonality

When occurring in a position of closure, the tonic is stable and evokes a pleasant experience. (So too, but to a lesser degree, do the mediant and dominant pitches.) Whatever

else one may say, the tonic is a familiar pitch at the ends of musical passages.

**One way to measure the similarity of fit is **to calculate the coefficient of correlation. **For a large sample of music, Huron (1992) **found that the average correlation between the **Krumhansl and Kessler key-profiles and the **frequency of occurrence of the scale degrees **was +0.88.

**Perhaps the most important observation to **be made about scale degree is that listeners **readily distinguish between major and minor **key contexts. **Krumhansl and Kessler's work implies that listeners **are readily able to switch their expectations depending **upon the modal context. **That is, listeners know to apply different **expectations for music depending on whether the **mode is major or minor.

Scale Degree Distributions and Tonality

When listeners rate the *stability* of various scale tones, they effectively replicate the frequency of occurrence of these tones in real music. This relationship strongly suggests that listeners experience the most commonly occurring tones as the most stable.

The word "tonality" is used by musicians in at least ten definable senses. One of the most common definitions of tonality is as a system of relating pitches or chords to some focal point or center -- the *tonic.* In Western music, these relationships are typically identified using scale-degree terms, such as tonic, supertonic, mediant, etc. Each of these scale-degrees evokes a different psychological

quality or character according to how it is heard in relation to the prevailing tonal center. As we saw earlier, by an act of will, musicians can imagine a single tone as either the leading-tone, mediant, or tonic, etc. The ability of listeners to imagine tones or chords as serving different tonal functions testifies to the cognitive (rather than perceptual) basis of tonality.

How does the tonic pitch become an internalized reference for listeners? The work of Carol Krumhansl suggests that tonal schemas are learned through exposure to music from a given culture or genre. Moreover, Krumhansl's work suggests that one of the primary factors influencing tonality perception is the simple frequency of occurrence of different tones. The most frequent pitch has a tendency to be heard as the tonic:

> "Listeners appear to be very sensitive to the frequency with which the various elements [pitch chromas] and their successive combinations are employed in music. It seems probable, then, that abstract tonal and harmonic relations are learned through internalizing distribution properties characteristic of the style." (Krumhansl, 1990; p.286).